

UNIVERSITÉ DU QUÉBEC

MÉMOIRE DE RECHERCHE PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN BIOPHYSIQUE

PAR
ALAIN DOUCET, B.Sc. (Biochimie)

COMBINAISON DE DEUX TYPES DE MUTATIONS STABILISANTES AU SEIN
DE LA PROTÉINE *DE NOVO* MB-1

JANVIER 2002

2091

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

Avant propos et remerciement

Ma soif de découvrir et d'apprendre comment fonctionne la nature et plus précisément les êtres vivants m'a poussée à entreprendre des études en sciences naturelles. Le baccalauréat en biochimie de l'U.Q.T.R. était le programme tout désigné pour acquérir les connaissances théoriques ainsi que l'aspect pratique de l'étude des composantes du vivant. J'ai donc entrepris et complété ce programme d'étude. Avant la rencontre du docteur Beauregard à ma dernière année du baccalauréat en biochimie je ne savais pas si je devais continuer mes études au cycle supérieur ou me diriger vers le marché du travail. Cette rencontre des plus intéressante a suscité en moi le désir d'apprendre le fonctionnement d'un laboratoire de recherche et la démarche scientifique s'y rattachant. Le sujet de travail et les techniques utilisées dans le laboratoire de M. Beauregard m'apparurent des plus intéressantes et je lui ai donc proposé de compléter un projet de fin d'étude au sein de son laboratoire. Le docteur Beauregard a bien voulu me donner ma chance et il a donc accepté de me superviser pour ce projet. À la fin de ce projet ma décision était prise : je désirais continuer mes études au cycle supérieur sous la direction de M. Beauregard. Il a donc accepté de m'intégrer à son équipe de recherche et m'a suggéré un projet de recherche des plus stimulant et intéressant. Je le remercie sincèrement car sans cette opportunité je n'aurais probablement pas continué mes études. Je le remercie aussi de m'avoir dirigé, supporté et encouragé pendant ces deux années d'études. C'est grâce à ce professeur si aujourd'hui j'aime la recherche scientifique et désire poursuivre dans cette voie.

L'équipe de recherche du laboratoire de M. Beauregard a aussi été très importante dans mon cheminement. Je tiens à remercier Martin Williams, Mylène Gagnon et Maxime Sasseville qui m'ont permis d'apprendre les techniques utilisées dans le laboratoire. Ils ont aussi créé un environnement propice au développement de nouvelles idées et ont aussi participé à des discussions très constructives.

Je tiens aussi à souligner l'importance des autres enseignants du programme de maîtrise en biophysique qui ont contribué activement à l'acquisition de nouvelles connaissances. Des professeurs et collaborateurs tels que Marc Sirois, François Boucher et Maria Grazia-Martinoli qui ont permis à mon projet d'avancer plus rapidement. Je remercie aussi tous les autres professeurs qui ont contribué à mon développement.

Une note toute spéciale pour mes parents et ma copine, Véronique Lafrance, pour leur appui et leur soutien moral. Ces personnes m'ont poussé à persévérer tout au long de mes études et je les en remercie.

Finalement, ces deux années d'études n'auraient pas été aussi enrichissantes sans le développement de relations sincères avec les autres étudiants gradués. Je pense entre autre à Keith Chiasson et Barbara Valastro avec qui j'ai pratiqué plusieurs activités sportives et eu des discussions très intéressantes. D'autres étudiants tels que Sylvie Gélinas, Marc-André Laurin et tous les étudiants ayant fait un stage dans notre laboratoire ont également égayé ces journées de manipulations.

RÉSUMÉ

La protéine MB-1 est une protéine à haute valeur nutritive destinée à l'alimentation animale. Cette protéine est aussi utilisée pour des études fondamentales visant à déterminer la relation entre la séquence et la structure des protéine hélicales. Pour pouvoir être utilisée dans l'alimentation animale, MB-1 se doit d'être stable. Pour l'instant cette protéine manque de stabilité et deux stratégies ont été mises au point afin de résoudre ce problème.

La première consiste en l'insertion de ponts disulfures chez la protéine. Deux variants de MB-1 ont été créés. Ces variants se démarquent par la position relative de leurs hélices. Cette différence de position a donnée naissance à la protéine de connectivité main droite et à celle de connectivité main gauche. L'ajout de ponts disulfures a amené une augmentation de stabilité thermique et protéolytique chez les protéines dont le pont disulfure est fermé. L'ouverture des ponts engendre cependant le dépliement des protéines indiquant que les mutations sont déstabilisantes tandis que la formation du pont est, quant à elle, stabilisante. La deuxième stratégie consiste à remplacer la tyrosine 62 par un tryptophane au sein de MB-1, créant ainsi la protéine MB-1Trp. Ce remplacement a mené à une grande augmentation de la stabilité de la protéine.

Le but de mon projet était de combiner les deux stratégies au sein d'une même protéine. J'ai donc procédé à des manipulations génétiques pour créer deux protéines

comportant chacune un pont disulfure et un tryptophane en position 62. J'ai exprimé ces protéines au sein de bactéries pour ensuite les isoler et les caractériser.

Les expériences de caractérisation démontrent que les protéines ont un haut contenu en structures secondaires hélicales. Leur cœur hydrophobe est formé mais il est très fluide. La protéine de connectivité main gauche conserve en partie ses structures hélicales lors du bris du pont disulfure. Une augmentation de la stabilité conformationnelle a donc été obtenue par l'insertion du tryptophane au sein de la protéine à pont disulfure de connectivité main gauche. Les deux mutants que j'ai produits sont aussi plus stables que la protéine MB-1 mais ne surpassent pas la protéine MB-1Trp. L'insertion du tryptophane au sein des protéines à ponts disulfures n'a pas amené un aussi grand effet stabilisant qu'escompté. Les modifications effectuées nous amènent cependant de l'information sur la relation qui existe entre la séquence de la protéine et la structure de celle-ci.

TABLE DES MATIÈRES

AVANT-PROPOS.....	i
RÉSUMÉ.....	ii
TABLE DES MATIÈRES.....	iv
LISTE DES TABLEAUX.....	vii
LISTE DES FIGURES.....	viii
LISTE DES SYMBOLES ET ABRÉVIATIONS.....	x
 CHAPITRE I INTRODUCTION.....	 1
1.1 Structures et stabilité des protéines globulaires.....	1
1.1.1 Niveaux structuraux présents au sein des protéines.....	1
1.1.2 Principes de stabilité d'une protéine globulaire.....	8
1.1.3 Les caractéristiques d'une protéine native.....	9
1.2 Les acides aminés essentiels et la vache laitière.....	10
1.2.1 Problèmes chez la vache laitière.....	10
1.2.2 Solutions possibles pour résoudre la carence en acides aminés essentiels.....	11
1.3 Création de la protéine MB-1 et de dérivés de cette protéine.....	15
1.3.1 Défauts de la protéine MB-1.....	19
1.3.2 Création de mutants à ponts disulfures.....	19
1.3.3 Création de MB-1Trp.....	23

1.4 Mon projet de maîtrise : combinaison des deux types de mutations au sein de MB-1	25
1.5 Présentation de l'article scientifique.....	26
CHAPITRE II RÉSULTATS.....	27
1.6 Article scientifique.....	28
Abstract.....	29
Introduction.....	30
Experimental Procedures.....	34
Results.....	41
Discussion.....	47
Abbreviations used.....	49
Acknowledgments.....	49
Literature Cited.....	50
Figure Captions.....	58
Figures.....	61
1.7 Tableau comparatif.....	67
CHAPITRE III DISCUSSION.....	68
1.8 Remise en contexte de notre travail.....	68
1.8.1 Type de protéine produite et méthode de production.....	68
1.9 Élément original de mon projet de recherche.....	70
1.10 Discussion des résultats obtenus.....	71
1.10.1 Les mutants TrpLH et TrpRH vs. la protéine MB-1.....	71
1.10.2 Insertion de ponts disulfures chez MB-1Trp.....	72
1.10.3 Ajout d'un tryptophane à la protéine MB-1RH et MB-1LH.....	74

1.10.4 MB-1TrpLH vs. MB-1TrpRH.....	76
CHAPITRE IV CONCLUSION.....	78
CHAPITRE V OUVERTURE DE RECHERCHES FUTURES.....	81
1.11 Spécification de la connectivité.....	81
1.12 Évolution accélérée et sélection à haut volume.....	82
1.13 Cristallographie et modélisation.....	83
BIBLIOGRAPHIE.....	84
ANNEXE A.....	87

LISTE DES TABLEAUX

1	Comparison of physical properties measured for the four mutants.....	60
2.1	Tableau comparatif des propriétés physiques de MB-1 et de différents mutants.	67

LISTE DES FIGURES

CHAPITRE I INTRODUCTION

1.1	Structure générale des acides aminés.....	3
1.2	Enchaînement de deux acides aminés à l'aide d'un lien peptidique.....	4
1.3	Principales structures secondaires présentes chez les protéines.....	5
1.4	Exemple de structure tridimensionnelle d'une protéine.....	7
1.5	Schéma des différents estomacs de la vache.....	12
1.6	Dégradation ruminale des protéines.....	13
1.7	Vue de haut d'un carré à quatre hélices.....	16
1.8	Représentation ouverte et aplatie de la protéine MB-1.....	17
1.9	Modèle à ruban d'une protéine à nœud de quatre hélices.....	18
1.10	Modèle à ruban des deux topologies que peut adopter un carré d'hélice.....	20
1.11	Structures chimiques des chaînes latérales de la méthionine, leucine et de la cystéine.....	22
1.12	Structure chimique de la tyrosine et du tryptophane.....	24

CHAPITRE II RÉSULTATS

1	Schematic representation of MB-1LH and MB-1RH.....	61
2	MB-1's hydrophobic core as predicted by design.....	62

3	CD spectra of the four mutants.....	63
4	CD spectra measured in the presence of DTT.....	64
5	Thermal denaturation curves.....	65
6	Proteolytic degradation of the proteins.....	66

LISTE DES SYMBOLES ET ABRÉVIATIONS

A, Ala	alanine
A.A.	acide aminé
A.A.E.	acide aminé essentiel
ANSA	8-anilino-1-naphtalenesulfonic acid
BCA	bicinconinique acid (acide bicinconinique)
C, Cys	cystéine
CD	circular dichroism (dichroïsme circulaire)
D, Asp	acide aspartique
DTT	dithiothreitol
E, Glu	acide glutamique
F, Phe	phenylalanine
H, His	histidine
K, Lys	lysine
L, Leu	leucine
LH	left handed (main gauche)
M, Met	méthionine
MB-1	Milk Bundle 1
N, Arg	arginine
P, Pro	proline
p-HMB	<i>para</i> -hydroxymercuribenzoate
Q, Gln	glutamine
RH	right handed (main droite)
S, Ser	sérine
SDS-PAGE	sodium dodecyl sulfate-polyacrilamide gel electrophoresis (gel électrophorétique de polyacrylamide contenant du dodécyl sulfate de sodium)
SEC	size exclusion chromatography (chromatographie d'exclusion de taille)
T, Thr	thrénine
T _m	température de dénaturation
UV/Vis	ultraviolet/visible
W, Trp	tryptophane
Y, Tyr	tyrosine

CHAPITRE 1

INTRODUCTION

1.1 Structures et stabilité des protéines globulaires

Le rôle d'un biochimiste et d'un biophysicien opérant dans le domaine de la recherche scientifique est de déterminer les lois naturelles régissant les organismes vivants. Le scientifique explore donc les différentes fonctions et molécules de ces organismes. Au sein de ceux-ci on retrouve plusieurs types de molécules d'intérêt tels les lipides, les glucides et les protéines. Les protéines m'intéressent particulièrement car ce sont des molécules avec des rôles physiologiques importants. Ces molécules présentent aussi plusieurs niveaux de structures qui leurs confèrent un arrangement tridimensionnel précis et pour la plupart, une activité biologique. Il est aussi connu que plusieurs maladies telle que l'Alzheimer et la thalassémie impliquent un mauvais repliement d'une ou plusieurs protéines. L'étude de la relation entre les différentes structures retrouvées chez une protéine et l'effet de ces structures sur les caractéristiques physico-chimique de la protéine présentent donc un intérêt certain.

1.1.1 Niveaux structuraux présents au sein des protéines

Avant d'aborder les protéines il faut d'abord connaître les sous-unités de base composants les protéines, soit les acides aminés (A.A.). L'enchaînement de plusieurs acides aminés forme une protéine (1). Comme le démontre la figure 1.1, un acide aminé est composé de quatre groupements chimiques ainsi que d'un atome d'hydrogène (2). Les quatre groupements sont :

1. Un groupement carboxyle conférant un caractère acide à la molécule.
2. Un groupement amine ayant la propriété d'être basique.
3. Un carbone α central.
4. Une chaîne latérale R rattaché au carbone alpha.

Il faut aussi savoir que l'on retrouve au sein des protéines 20 différents acides aminés principaux qui se distinguent par la nature de leur chaîne latérale. Ces chaînes latérales contiennent différents groupements chimiques leur conférant différentes propriétés physico-chimiques. Certaines chaînes sont hydrophobes alors que d'autres sont polaires. Certaines sont aussi chargées à pH physiologique.

Ces acides aminés peuvent être rattachés les uns aux autres comme le démontre la figure 1.2 (3). L'ordre d'enchaînement de plusieurs acides aminés forme la structure primaire d'une protéine. C'est aussi la séquence dans laquelle les acides aminés sont placés qui déterminera les autres niveaux de structures de la protéine et les caractéristiques physiques et chimiques de cette même protéine.

Les acides aminés attachés les uns aux autres peuvent se positionner de plusieurs façon dans l'espace pour former des structures périodiques appelées structures secondaires des protéines. Nous retrouvons principalement deux types de structures périodiques au sein des protéines soit les hélices α et les feuillets β (4). Ces deux types de structures sont présentés à la figure 1.3. Dans la partie du haut de cette figure nous voyons la structure chimique des acides aminés formant les deux types de structures alors qu'au bas de la figure nous avons la représentation en ruban du squelette des différentes structures secondaires. Comme on peut le constater, ce qui caractérise les structures secondaires est la position dans l'espace des groupements de chaque côté du carbone alpha. La répétition des angles dièdres pour chacun des acide aminé de la

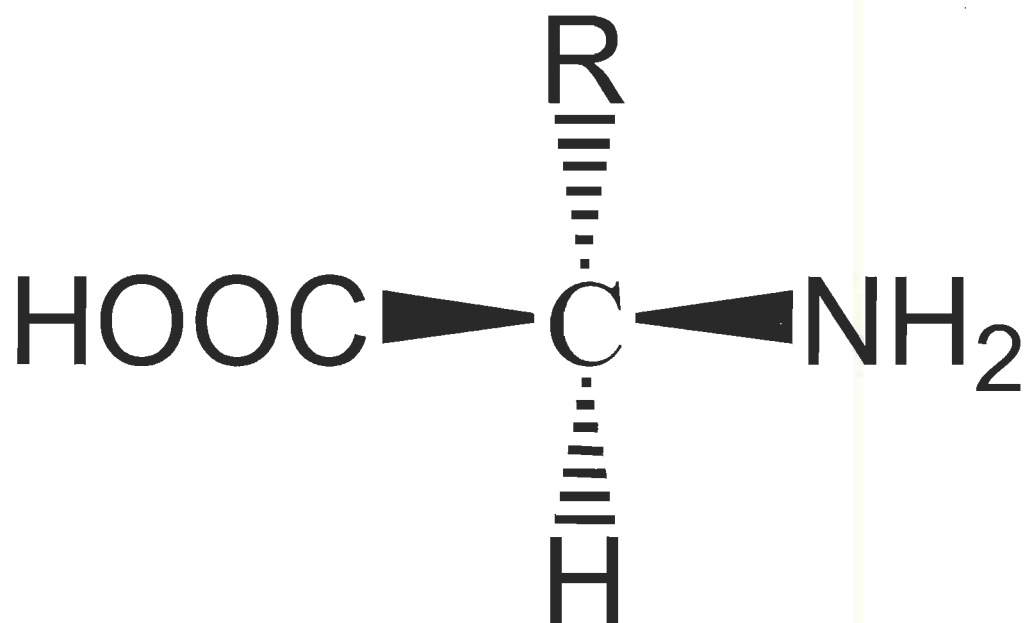


FIGURE 1.1. Structure générale des acides aminés. Le carbone central est appelé carbone alpha, le groupement amine se retrouve à droite de la figure alors que le groupement carboxyle se retrouve à gauche. La chaîne radicalaire est représentée par le R.

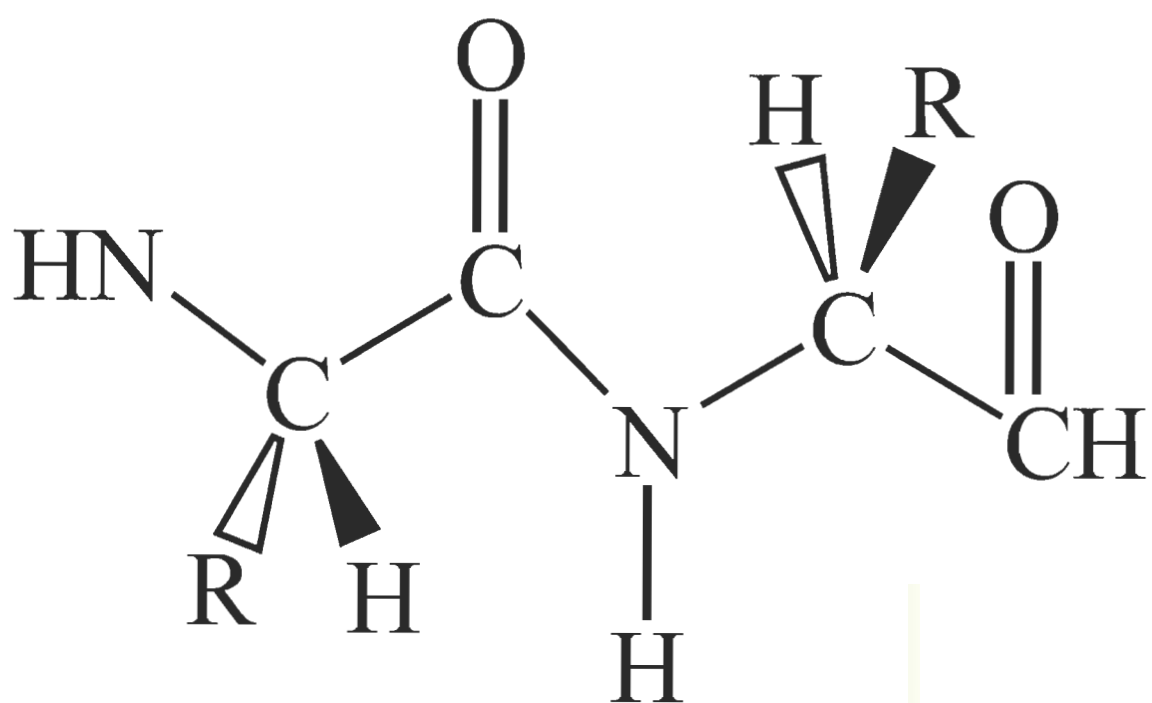


FIGURE 1.2. Enchaînement de deux acides aminés à l'aide d'un lien peptidique.

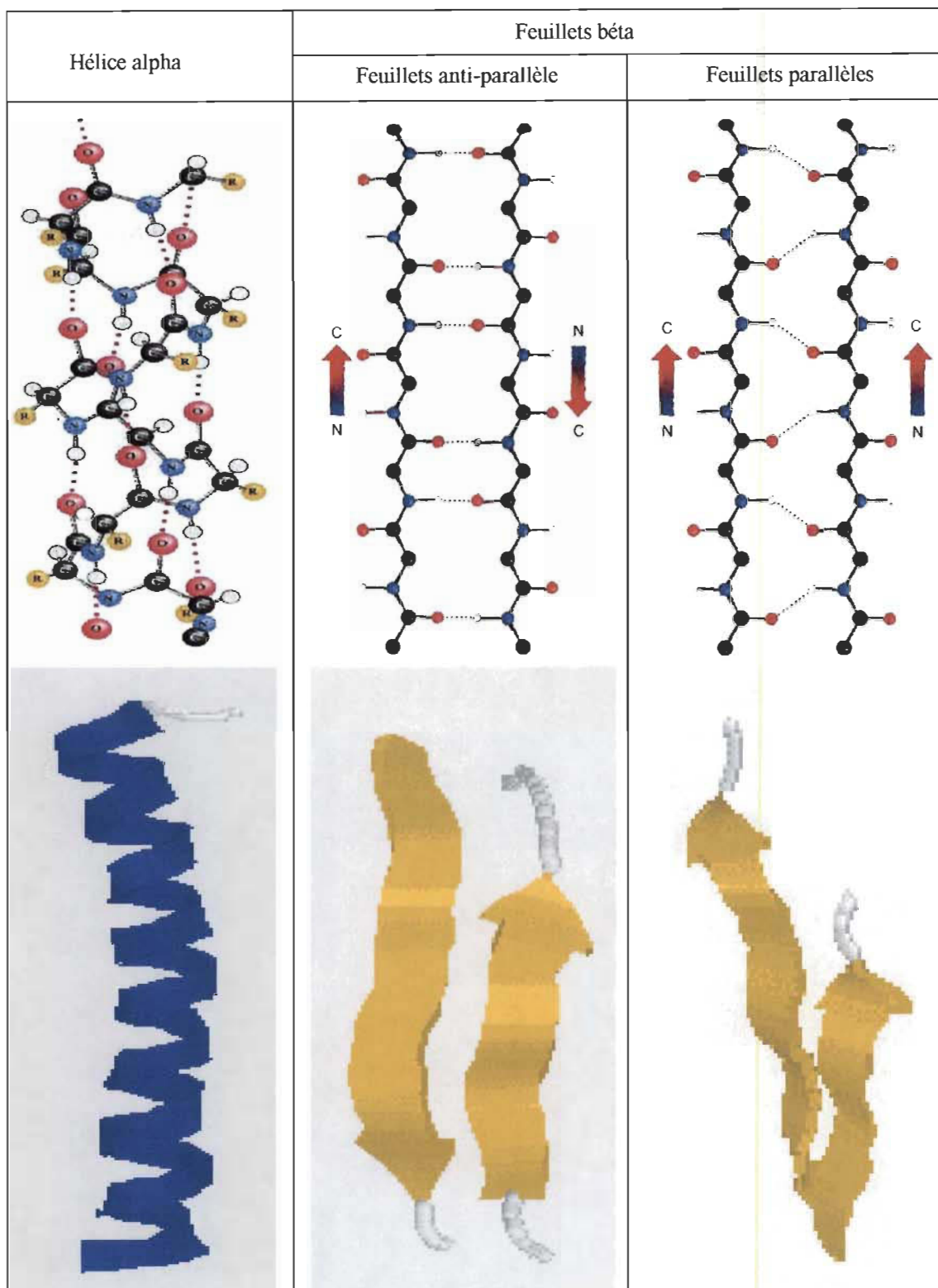


FIGURE 1.3. Principales structures secondaires présentes chez les protéines. Le haut de la figure présente les structures chimiques alors que le bas présente des modèles à ruban de ces mêmes structures.

chaîne amène la formation de structures périodiques. Le réseau de ponts hydrogènes créé au sein de la structure est aussi une caractéristique importante de celle-ci puisque ce réseau est très différent pour les hélices α et les feuillets β comme le démontrent les traits pointillés à la figure 1.3.

Lorsque les structures secondaires sont formées, elles peuvent se replier pour former la structure tertiaire de la protéine. La figure 1.4 présente un exemple de structure tertiaire (4). Il faut aussi savoir que la force majeure qui pousse une protéine globulaire à former une structure tertiaire est la force hydrophobe. C'est-à-dire le besoin de cacher les groupements non polaires de l'eau environnante. Les différentes structures secondaires s'associeront donc de façon à emballer les groupements hydrophobes dans le centre de la protéine et ainsi les isoler de l'eau (5). Si la protéine n'arrive pas à former ces structures il y aura instabilité de la protéine et celle-ci ne pourra pas remplir les fonctions pour lesquelles elle a été produite.

Les protéines sont donc des molécules adoptant des structures particulières et précises. On les retrouve au sein de tous les organismes vivants et elles sont essentielles à la vie. Elles remplissent de multiples fonctions telles que :

1. La catalyse de réactions chimiques en ce qui concerne les enzymes.
2. Un rôle structural pour le maintien de la forme des cellules ainsi que pour leur adhésion. Nous pouvons nommer la protéine tubuline qui compose les microtubules tandis que la famille des cadhérines a comme fonction d'attacher les cellules les unes aux autres.
3. Contribution à la motilité cellulaire. Pensons à l'actine et la myosine responsable de la contraction musculaire.

Toutes ces fonctions sont possibles grâce au repliement en trois dimensions des protéines. Il est donc d'une importance cruciale de comprendre et maîtriser les règles régissant le repliement de celles-ci.

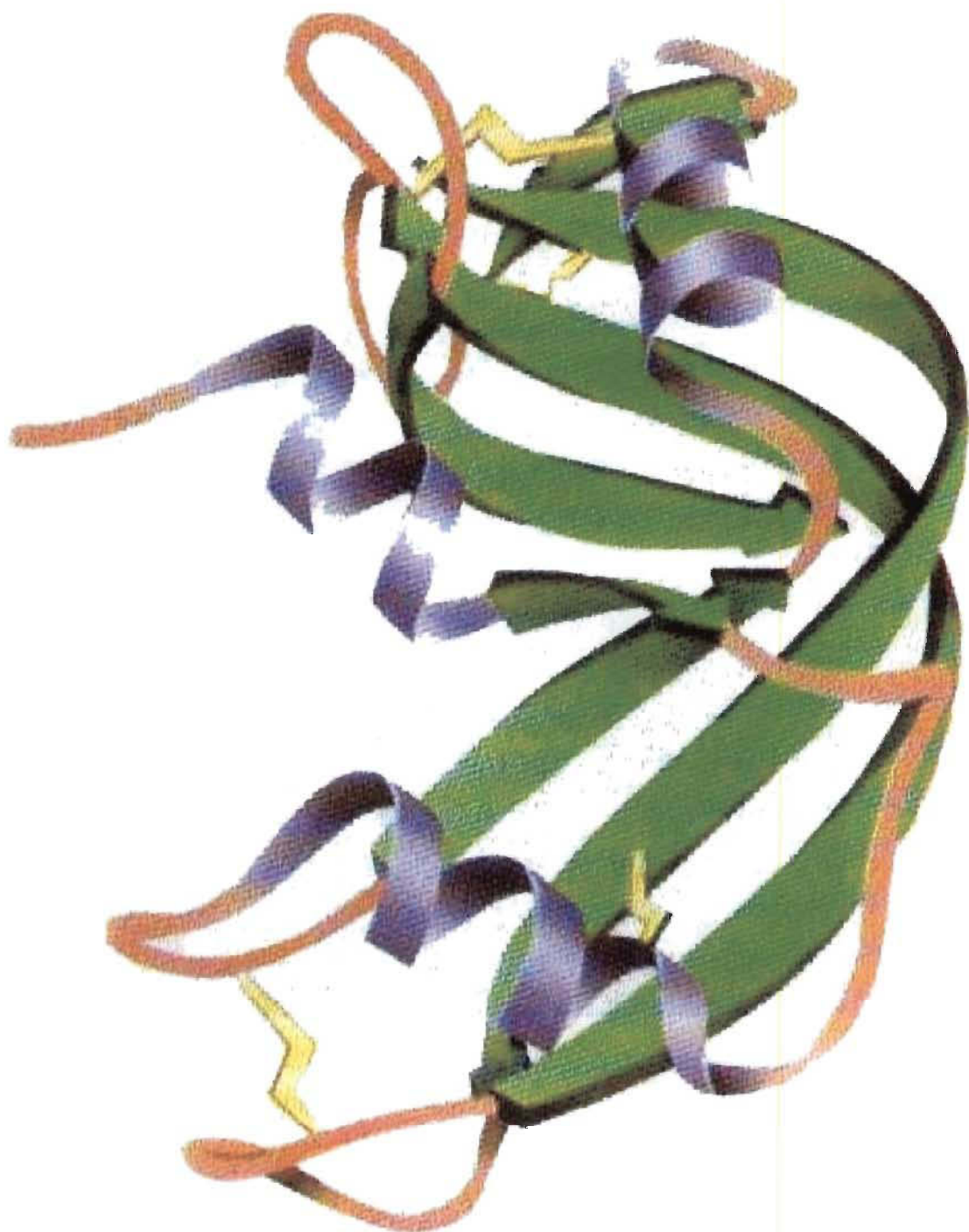


FIGURE 1.4. Exemple de structure tridimensionnelle d'une protéine. C'est le repliement précis des structures secondaires l'une par rapport à l'autre qui donne la structure tertiaire de la protéine. L'exemple présenté est la structure de la ribonucléase A bovine.

1.1.2 Principe de stabilité d'une protéine globulaire

La stabilité d'une protéine est mesurée par la différence d'énergie libre du système (ΔG), différence d'énergie mesurée entre la forme pliée et dépliée de la protéine à une certaine température (6). Les lois de la thermodynamique stipulent que la différence d'énergie libre est en relation directe avec la différence d'entropie (ΔS) et d'enthalpie (ΔH) des deux états du système (protéine pliée, dépliée) selon l'équation mathématique :

$$\Delta G = \Delta H - T\Delta S \quad [1.1]$$

Où T est la température en degré Kelvin.

Le terme enthalpique (ΔH) de l'équation est en relation avec des forces physiques telles que les interactions de Van der Waals, la formation de pont hydrogènes et les interactions de type électrostatiques chez les protéines (7). La force, la fréquence et la nature de ces interactions changent de façon drastique selon que la protéine se retrouve sous forme pliée ou dépliée. Ainsi, chez une protéine dépliée, la majorité des ponts hydrogènes sont formés entre les molécules d'eau et la protéine alors que pour une protéine bien repliée la majorité des ponts sont formés entre les résidus de la protéine.

Le terme entropique (ΔS) fait, quant à lui, référence à la liberté de mouvement des molécules du système. La deuxième loi de la thermodynamique stipule que l'univers tend à être de plus en plus libre et en désordre. Plus le désordre augmente et plus l'entropie augmente. Ainsi dans un système donné maintenu à volume et température constante tout phénomène spontané du système amènera une augmentation de l'entropie (8). Comment ce concept peut-il être appliqué à la stabilité d'une protéine ? Eh bien, nous savons que la solvation de molécules hydrophobes dans l'eau force l'eau à se structurer et à s'organiser, ce qui diminue l'entropie globale du système. Ceci est donc très défavorable. La seule façon d'empêcher ce phénomène pour les protéines globulaires est de compacter tous ses résidus hydrophobes au cœur de la protéine et ainsi

de les isoler de l'eau. Ce phénomène appelé force hydrophobe est la force majeure qui amène une protéine à se replier à une certaine température (7). À température physiologique (37°C) la majorité des protéines naturelles globulaires se retrouvent sous forme pliée.

Un phénomène physique sera spontané si l'énergie libre du système est diminuée. Cette énergie libre étant en relation directe avec le changement d'enthalpie et d'entropie entre les deux états du système comme mentionné à l'équation 1.1. Une protéine se repliera ou se dépliera donc spontanément à une température donnée de façon à atteindre l'état dans lequel son énergie libre sera minimale pour cette température (7).

1.1.3 Les caractéristiques d'une protéine native

La grande majorité des protéines naturelles possèdent les caractéristiques d'une protéine native. Pour être considérée dans un état natif une protéine globulaire doit rencontrer un certain nombre de critères. Premièrement, ses structures secondaires doivent toutes être formées. Deuxièmement, les structures secondaires doivent être positionnées de façon à bien cacher les résidus hydrophobes au centre de la protéine formant ainsi le cœur hydrophobe de la protéine. Les résidus se trouvant dans le cœur hydrophobe doivent se retrouver dans un état quasi cristallin. Ce qui signifie que leur position est fixe et rigide et que le même résidu se retrouve exactement dans la même conformation chez plus de 99% des protéines présentes en solution. Le noyau hydrophobe d'une telle protéine est donc très peu fluide. En ce qui concerne la stabilité thermodynamique, les petites protéines globulaires à l'état natif ont une dénaturation de type coopérative. Ce qui signifie que la protéine passe de la forme pliée à la forme dépliée sur une très petite gamme de température lors de son chauffage. En ce qui concerne les protéines globulaires de taille importante, elles présentent généralement plusieurs domaines distincts et c'est la dénaturation de chaque domaine qui est de type

coopératif. Finalement, une protéine native est compacte, elle offre très peu de cibles aux protéases et résiste donc assez bien à celles-ci

1.2 Les acides aminés essentiels et la vache laitière

Nous savons donc que les protéines sont composées d'acides aminés. D'où proviennent donc ces molécules ? Les sources d'acides aminés sont multiples pour un organisme vivant. Plusieurs acides aminés peuvent être synthétisés par l'organisme lui-même à partir de métabolites provenant de glucides. Toutefois, la plupart des organismes sont incapables de produire tous les acides aminés à partir de leur seul métabolisme. Ces molécules sont nommées acides aminés essentiels (A.A.E.) et il faut donc à l'organisme une source externe pour ces acides aminés particuliers (9). En ce qui concerne les organismes supérieurs tels que les mammifères, la source d'A.A.E. externe est leur diète. Un animal doit pouvoir retrouver une quantité suffisante d'A.A.E. dans ses aliments pour être en mesure de synthétiser les protéines dont il a besoin.

1.2.1 Problèmes chez la vache laitière

La vache laitière présente un problème particulier en ce qui a trait à son apport en acides aminés essentiels. Une vache laitière, en plus d'avoir à produire les protéines nécessaires à sa croissance et sa survie, doit produire les protéines du lait en grande quantité. Cette situation exige donc un excellent apport en nutriments et plus particulièrement en A.A.E. Les méthodes traditionnelles d'élevage peuvent mener à une carence en A.A.E. (10, 11) chez les vaches laitières et ce pour deux raisons principales :

1. La diète de la vache est assez pauvre en acides aminés méthionine (M, Met), thréonine (T, Tyr), lysine (K, Lys) et leucine (L, Leu).
2. Il y a une bio transformation des acides aminés par les microbes contenus dans le rumen de la vache.

Le point 2 fait référence au rumen de la vache. Le rumen est un compartiment stomacal de la vache. La figure 1.5 présente les différents estomacs de la vache. Le rumen compte pour environ 80% du volume total des différents estomacs. Nous pouvons voir que le rumen est le premier compartiment du système digestif et que tous les aliments ingurgités par la vache y séjournent. Le pH du rumen avoisine la neutralité et il y règne une flore microbienne importante. Ce sont ces micro-organismes qui sont responsables de la biotransformation des acides aminés de la diète et donc responsable de la quantité d'A.A.E. qui sera disponible pour la vache. La figure 1.6 schématise le processus de dégradation ruminale. On constate que la majorité des protéines de la diète seront hydrolysées en acides aminés libres qui sont alors absorbés par les micro-organismes. Ceux-ci transforment les A.A. absorbés afin de produire les protéines microbiennes (12). Ce phénomène change le profil d'A.A.E. disponibles pour la vache. Ce changement en A.A.E. peut affecter le métabolisme des vaches et augmenter le coût de production du lait. Ainsi, de façon à contrer partiellement cette carence les agriculteurs sont dans l'obligation de fournir plus d'A.A.E., par exemple en augmentant le volume d'aliment fourni au bétail.

1.2.2 Solutions possibles pour résoudre la carence en acides aminés essentiels.

Pour résoudre le problème de carence en A.A.E. chez la vache laitière différentes solutions ont été proposées. La première solution consiste à supplémenter la nourriture fournie à l'animal à l'aide de capsules contenant des acides aminés. Ces capsules d'acides aminés sont disponibles commercialement sous le nom de SmartAminesTM et sont vendues par la compagnie Aventis (13). L'effet le plus marqué de ces capsules est l'augmentation de la quantité de protéines présentes dans le lait, ce qui augmente par le fait même la qualité du lait. Toutefois, cette solution présente des désavantages. Le premier est le coût élevé des capsules. Le deuxième provient de la difficulté à protéger adéquatement les acides aminés à l'aide d'une capsule lors de leur passage dans le rumen sans diminuer pour autant le relargage ultérieur de ces acides aminés une fois les

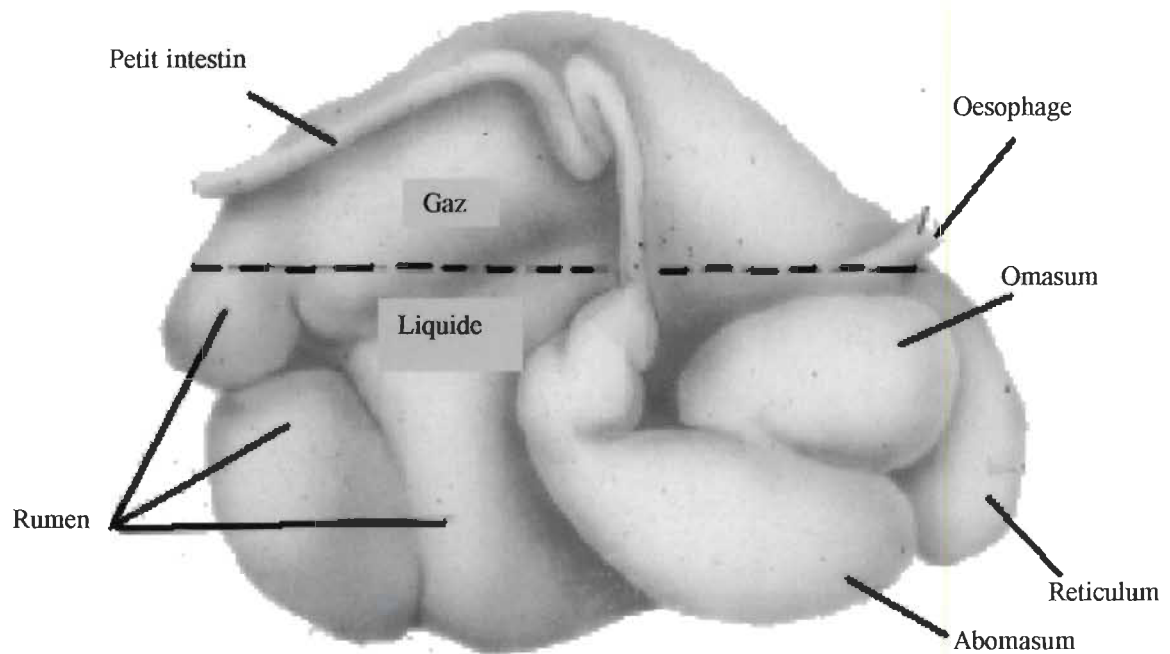


FIGURE 1.5. Schéma des différents estomacs de la vache. Le rumen compte pour environ 80% du volume total des estomacs.

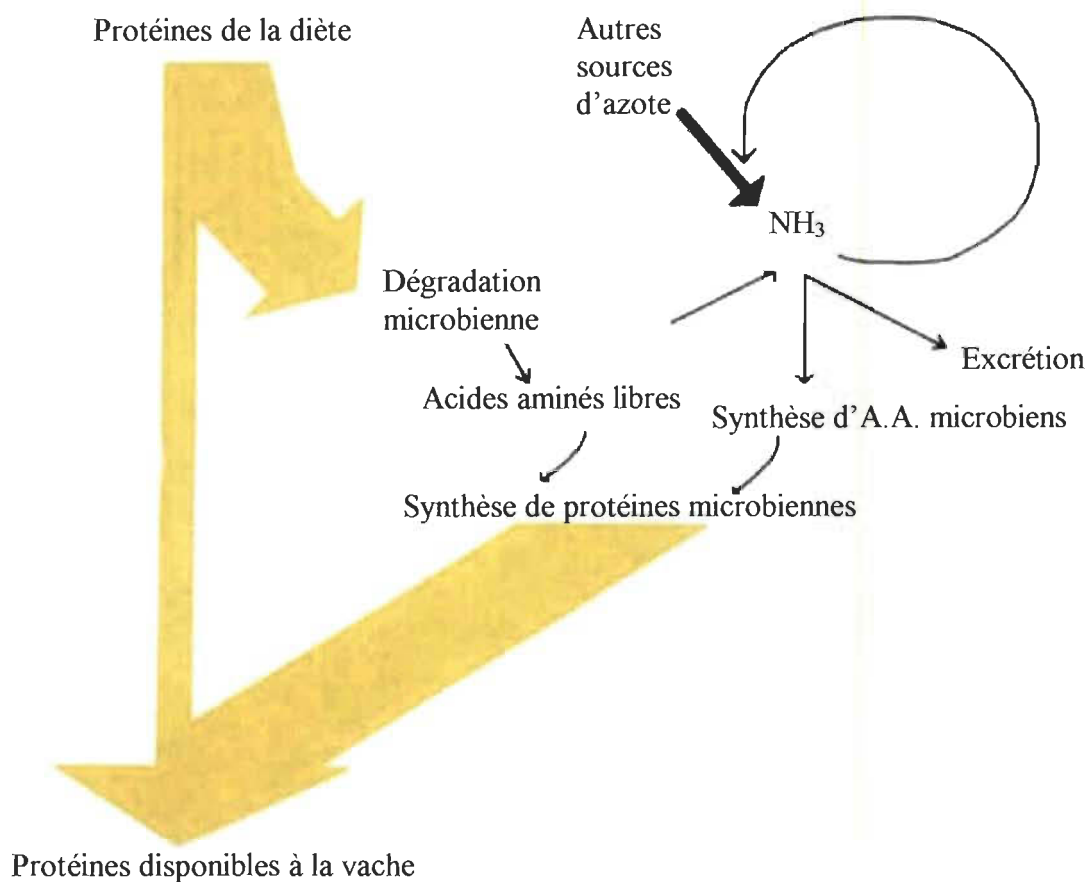


FIGURE 1.6. Dégradation ruminale des protéines. On constate que la majorité des protéines qui transitent dans le rumen sont digérées par les micro-organismes qui résident dans celui-ci.

capsules arrivées dans le petit intestin de l'animal.

Les deux autres solutions mentionnées dans ce texte consistent en la mise au point de protéines enrichies en A.A.E.. Ces protéines pourront ensuite être produites dans des plantes modifiées génétiquement entrant dans la diète de la vache par exemple. Pour que cette solution soit efficace il faudra que les protéines enrichies soient très stables de façon à pouvoir résister au maximum aux protéases ruminales.

La deuxième solution consiste à modifier une protéine naturelle de façon à pouvoir y incorporer un maximum d'A.A.E.. Peu de tentatives en ce sens ont été fructueuses jusqu'à maintenant. La raison principale de cet insuccès réside dans le fait que la modification d'un grand nombre de résidus au sein d'une protéine résulte habituellement en une baisse de stabilité de la protéine (14, 15). La protéine est dans l'incapacité de former les structures secondaires et tertiaires habituelles et ne se replie donc pas. Une protéine non repliée offre une multitude de cibles pour les protéases se retrouvant dans le rumen et serait donc rapidement dégradée. Une telle protéine serait donc inefficace pour contrer la carence en A.A.E. chez le bovin. Une autre difficulté rencontrée lors de la modification d'une protéine naturelle est le fait qu'il faille enrichir celle-ci de plusieurs types d'acides aminés en même temps pour qu'elle puisse répondre au problème de la carence. Ceci rend l'opération de modification très complexe et fastidieuse.

La troisième solution consiste en la création d'une toute nouvelle protéine adaptée au problème nutritionnel rencontré pour les vaches laitières. Cette solution a été choisie par notre laboratoire pour plusieurs raisons. La première étant que nous pourrions choisir la séquence d'acides aminés qui composera la nouvelle protéine et y incorporer un maximum d'A.A.E.. Nous pourrions spécifier la structure qu'adoptera la protéine en faisant le bon choix de séquence d'acides aminés. Finalement, cette protéine nous

permettrait d'étudier la relation entre la séquence d'une protéine et la structure qu'adoptera celle-ci. Pour que cette nouvelle protéine remplisse bien son rôle elle devra contenir un maximum d'A.A.E. choisis dans sa séquence et être résistante aux protéases et à la dénaturation thermique.

1.3 Création de la protéine MB-1 et de dérivés de cette protéine

L'équipe de recherche du Dr. Marc Beauregard créa en 1995 une toute nouvelle protéine appelée Milk Bundle-1 (MB-1) (16). La séquence de cette protéine a été choisie afin d'incorporer un maximum d'A.A.E. choisis. De plus, cette séquence est construite de façon à spécifier un repliement en nœud de quatre hélices pour la protéine MB-1. Afin de spécifier le type de repliement, le modèle de l'heptade a été utilisé pour choisir la séquence en acide aminé qui composera les hélices de la protéine. Un schéma de quatre heptades formant un carré d'hélice est présenté à la figure 1.7. Le modèle de l'heptade se base sur une définition de sept positions récurrentes pour compléter deux tours d'hélice. Afin de former un carré de quatre hélices, les positions « a » et « d » de l'heptade sont occupées par des résidus hydrophobes. La figure 1.8 présente la séquence ou structure primaire de la protéine MB-1, on y voit les quatre hélices de la protéine qui ont été aplaties dans le plan et elles sont représentées par les quatre parallélogrammes de couleurs. On remarque une forte présence des acides aminés M, T, K, L qui comptent pour 57% des résidus de la séquence. La figure 1.9 présente un modèle à ruban d'une protéine à nœud de quatre hélices. La structure de la protéine MB-1 devrait être semblable à celle de cette protéine si MB-1 se replie comme spécifié lors du design. Les résultats expérimentaux obtenus avec MB-1 démontrent que la protéine a un haut contenu en structures hélicales et que son cœur hydrophobe est bien formé. Malgré l'absence de structure cristallographique nous pouvons donc affirmer que la protéine se replie bien, les hélices sont formées et elles s'associent ensemble pour former une structure tertiaire compacte (16). Cependant, la protéine MB-1 ne possède pas que des qualités.

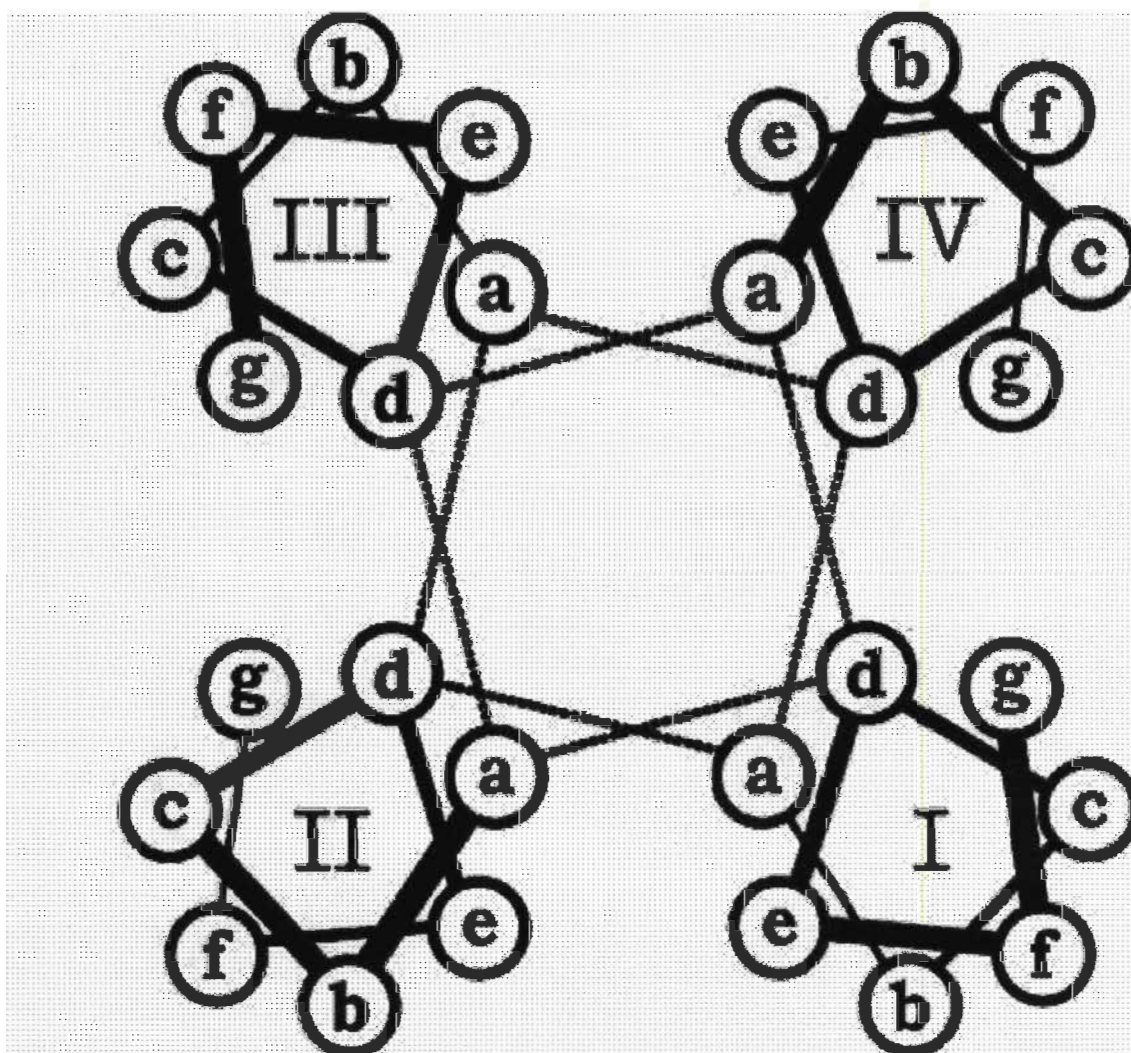


FIGURE 1.7. Vue de haut d'un carré à quatre hélices. Les sept positions récurrentes de l'heptade sont définies par les lettres de « a » à « g ». L'heptade définit les positions des résidus pour deux tours d'hélice.

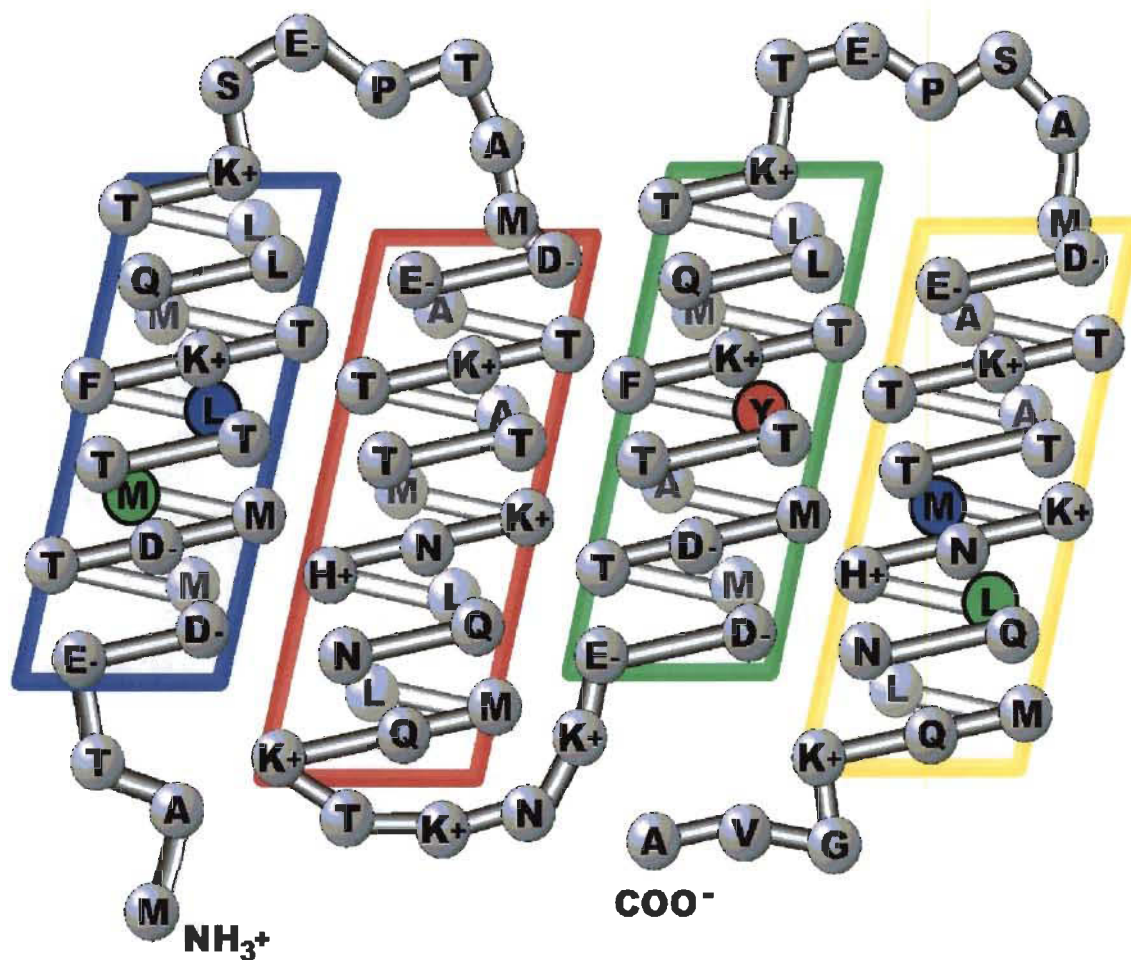


FIGURE 1.8. Représentation ouverte et aplatie de la protéine MB-1 démontrant la séquence en acides aminés de la protéine. Les parties hélicales de la protéine sont déterminées par les parallélépipèdes de couleur.

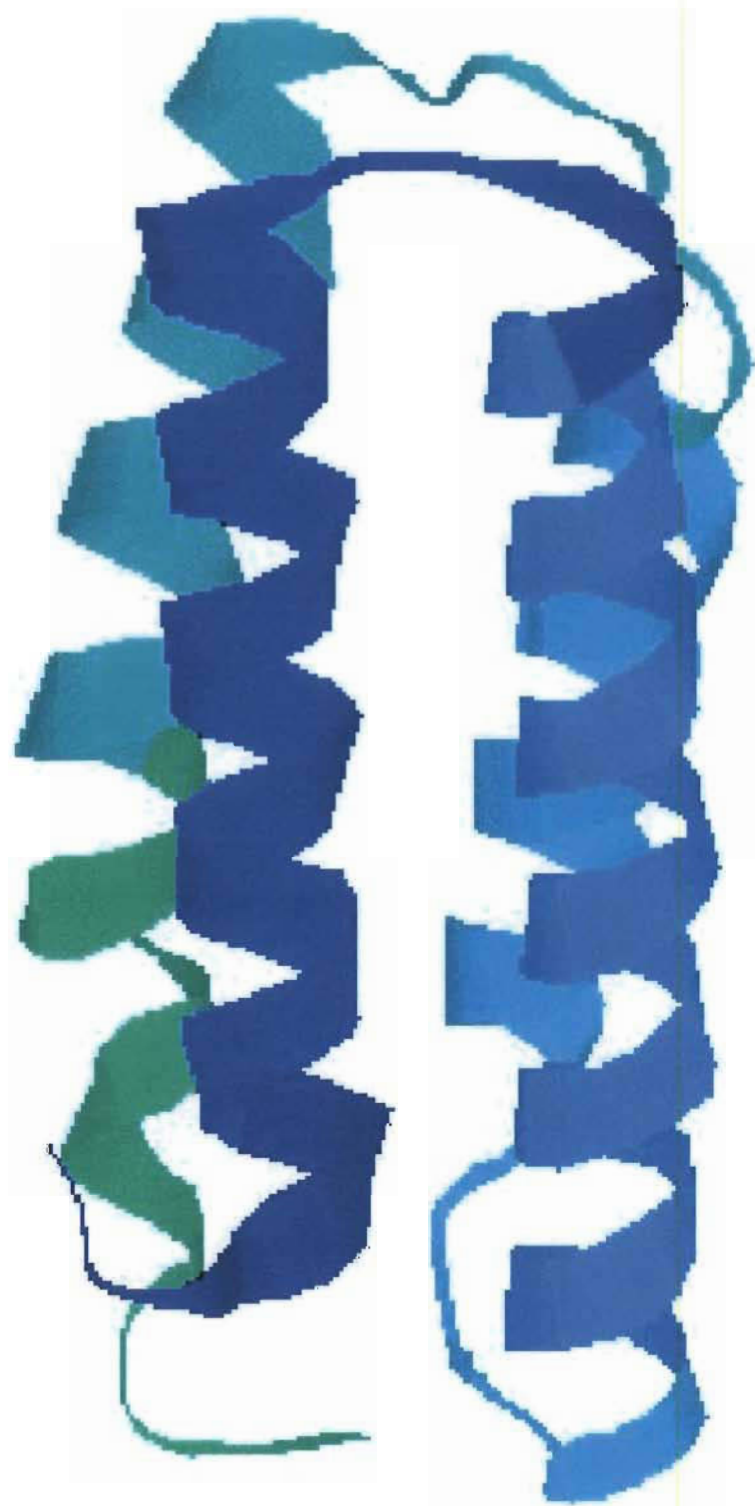


FIGURE 1.9. Modèle à ruban d'une protéine à nœud de quatre hélices.

1.3.1 Défauts de la protéine MB-1

La première ombre au tableau provient du fait que la protéine s'associe sous forme de dimère alors que l'on s'attend à la retrouver à l'état monomérique (17). Le deuxième défaut de MB-1 est sa faible résistance thermique et protéolytique. En effet, MB-1 a une température de dénaturation de 39⁰C et résiste très mal à la dégradation protéolytique. Si la protéine MB-1 était utilisée pour nourrir les vaches laitières elle serait rapidement dégradée dans le rumen et donc inefficace (18). Des solutions ont été proposées pour améliorer sa stabilité. Deux types de mutants ont été mis au point.

1.3.2 Création de mutants à ponts disulfures

Le premier type de variant à avoir été créé porterait un pont disulfure entre l'hélice 1 et 4 de la protéine MB-1. Deux buts étaient visés lors de cette modification :

1. Augmentation de la stabilité de la protéine MB-1. La formation d'un pont disulfure a en effet été prouvé efficace pour augmenter la stabilité thermique et protéolytique chez d'autres protéines. Son effet se répercute à plusieurs niveaux. Premièrement, il maintient rapproché plusieurs résidus augmentant les interactions de Van der Waals. Il réduit aussi beaucoup la liberté du squelette des deux hélices impliquées dans la formation du pont. L'impact de cette réduction d'entropie est majeur pour la protéine sous forme dépliée, ce qui favorise le repliement de la protéine (19).
2. Spécification de la topologie de la protéine MB-1. En effet, comme le démontre la figure 1.10, la protéine peut adopter deux structures différentes. La différence entre ces deux structures provient de la disposition des hélices l'une par rapport à l'autre. On constate que l'on peut replier la deuxième hélice à droite ou à gauche de la première et cette caractéristique spécifie la topologie main droite (RH) ou main gauche (LH) pour la protéine MB-1.

Très peu d'attention a été apportée à ce type de différence de topologie dans le domaine de l'ingénierie des protéines. La majorité des groupes de recherche travaillent avec des peptides formés de une à deux hélices qui s'associent en différents états d'oligomérisation. Comme les différentes hélices ne sont pas liées les unes aux autres il est impossible d'en déterminer la connectivité. Toutefois, lorsque l'on observe les protéines naturelles, la majorité des nœuds de quatre hélices adoptent la topologie main droite alors qu'une minorité de protéine, telle que la famille des ferritines, se replie selon une topologie main gauche. La séquence en acides aminés des protéines naturelles

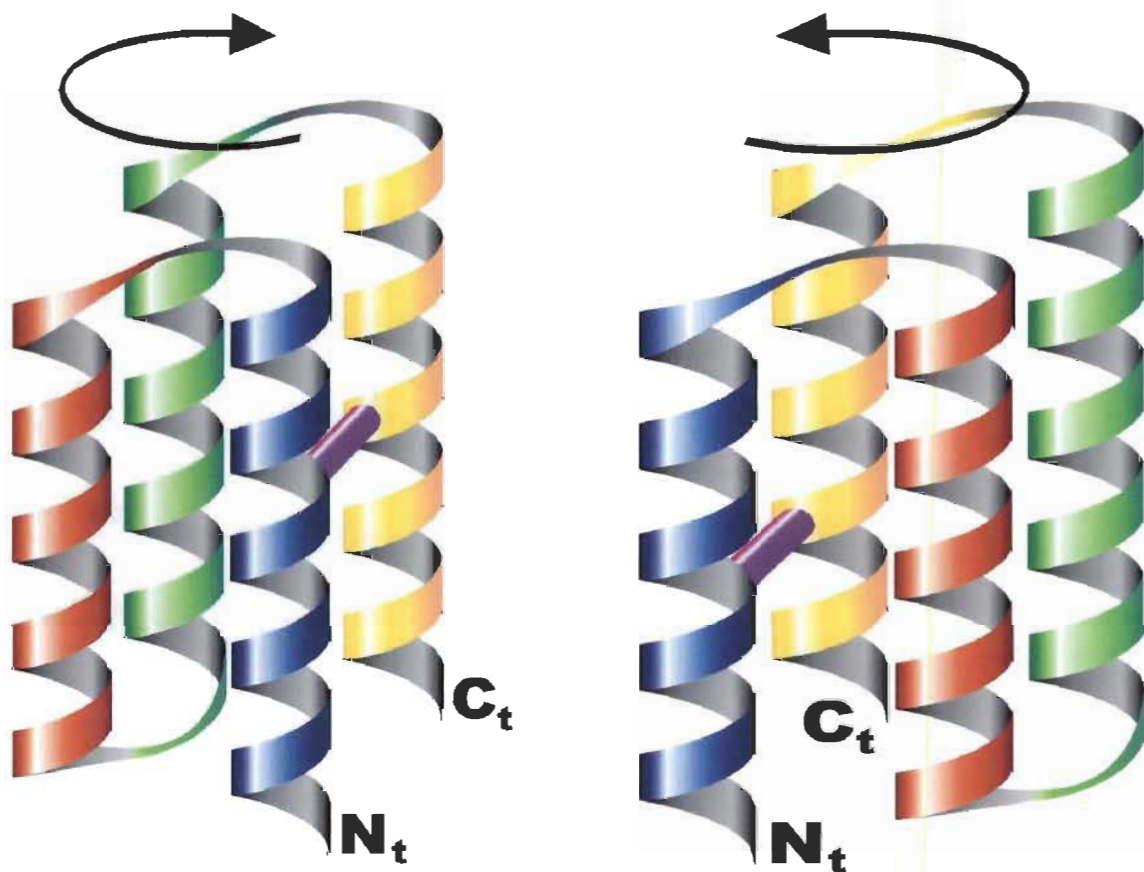


FIGURE 1.10. Modèle à ruban des deux topologies que peut adopter un carré d'hélice. La partie gauche démontre la topologie main gauche ou LH et la partie de droite, la topologie main droite ou RH. La différence de topologie provient de la position relative des hélices.

doit donc contenir l'information spécifiant le type de topologie que la protéine devra adopter. Des mutants de la protéine MB-1 ont été créés afin de spécifier la topologie de cette protéine.

Deux mutants comportant chacun un pont disulfure ont été créés. Un mutant, dont le pont spécifie la topologie main droite, a été nommé MB-1RH et un mutant de topologie main gauche nommé MB-1LH. Pour créer ces mutants la méthionine 10 et la leucine 91 ont chacune été remplacées par des cystéines pour former la protéine MB-1RH. Pour la formation du mutant MB-1LH nous avons remplacé la leucine 13 et la méthionine 87. La position des ponts est indiquée par un trait entre l'hélice 1 et 4 des protéines présentées à la figure 1.10.

L'insertion de ponts disulfures chez MB-1 a eu un effet stabilisant. Elle a amené l'augmentation de la température de dénaturation ainsi que l'augmentation de la résistance aux protéases et ce, pour les deux différents mutants (LH et RH). Toutefois, lorsque le pont disulfure est brisé à l'aide d'agent réducteur tel que le DTT nous observons une perte importante de structures hélicales chez les deux protéines à pont disulfure. Nous pouvons donc faire une distinction très nette entre l'effet de la formation d'un pont disulfure au sein de MB-1 et l'effet du remplacement de deux résidus du cœur hydrophobe par des cystéines. L'insertion des cystéines est catastrophique puisqu'elle induit un dépliement presque total de la protéine alors que la formation du pont est stabilisante. L'effet déstabilisant des cystéines est facile à comprendre lorsque l'on regarde la structure chimique des résidus impliqués lors de la mutation. La structure chimique des chaînes latérales de la méthionine, leucine et cystéine sont présentées à la figure 1.11. On constate que le volume de la chaîne latérale de la méthionine et de la leucine est plus important que celui de la cystéine. Ce qui indique que les mutations menant à la formation des protéines à pont disulfure ont induit une baisse du volume du cœur hydrophobe de la protéine et peut-être même créé une cavité, ce qui serait déstabilisant pour la protéine. Cet effet a aussi été observé pour la Rnase S où la

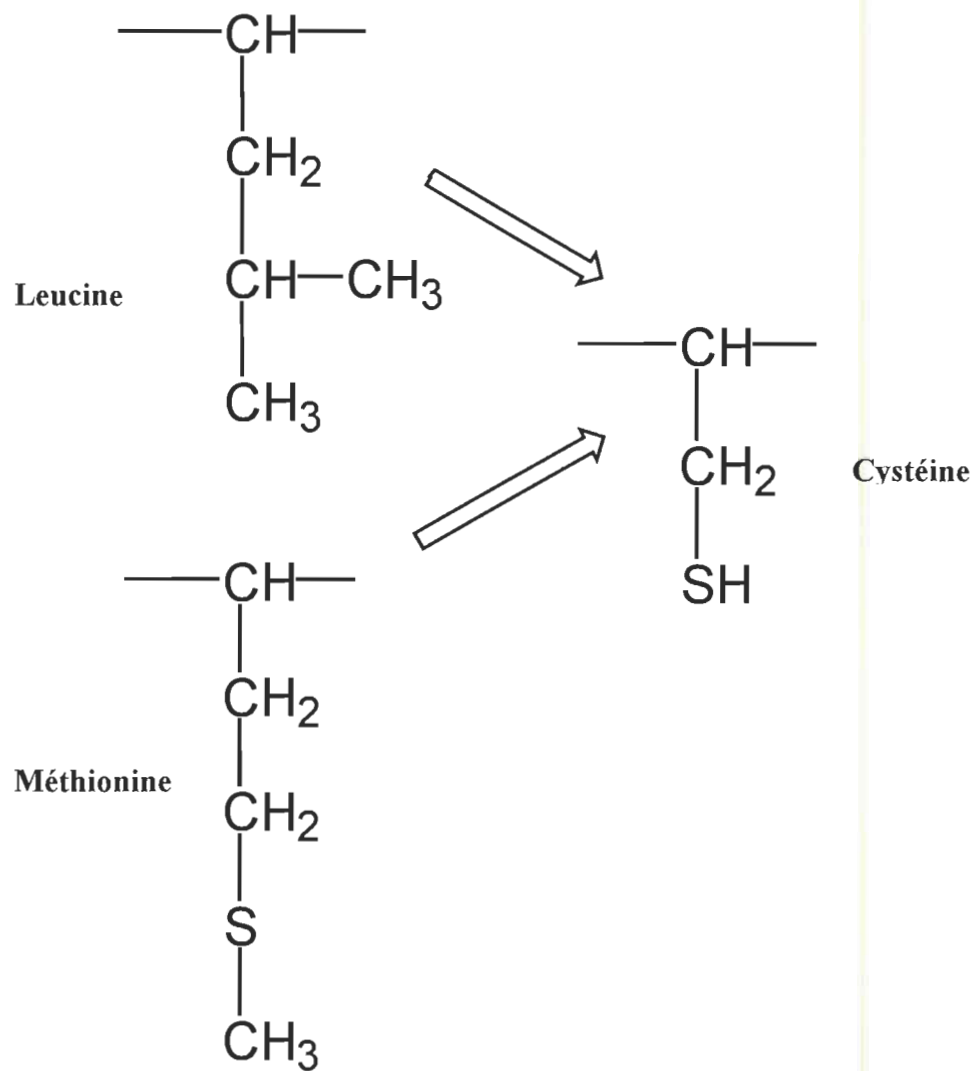


FIGURE 1.11. Structures chimiques des chaînes latérales de la méthionine, leucine et de la cystéine.

substitution d'une phénylalanine par des résidus plus petits amenait la formation de cavité et induisait une baisse de stabilité thermique pour la majorité des mutants (5). De plus, les résidus M et L sont beaucoup plus hydrophobes que la cystéine. Leur remplacement par la cystéine a donc diminué le caractère hydrophobe du cœur des protéines MB-1LH et MB-1RH et amené une déstabilisation des protéines. La formation du lien covalent entre les cystéines chez ces protéines réussit cependant à compenser pour l'effet déstabilisant des cystéines. L'effet global est donc une augmentation de la stabilité des protéines MB-1RH et MB-1LH par rapport à MB-1 lorsque leur pont disulfure est formé.

1.3.3 Création du mutant MB-1Trp

Un tout autre type de mutation a été effectué sur MB-1 dans le but d'en augmenter la stabilité. Pour la création du mutant MB-1Trp la tyrosine (Tyr) 62 de MB-1 a été remplacée par un tryptophane (Trp). Le but de cette mutation était d'augmenter le volume et l'hydrophobicité du cœur de la protéine. On voit très clairement à la figure 1.12 que la tyrosine ne comporte qu'un cycle aromatique alors que le tryptophane en présente deux. On constate aussi que la tyrosine présente un groupement hydroxyle polaire alors que le Trp n'en présente pas. L'effet du remplacement de la Tyr par un Trp chez MB-1 a été spectaculaire. La température de dénaturation est passée de 39⁰C à 55⁰C pour MB-1 et MB-1Trp, respectivement. Cet effet s'est aussi répercuté sur le degré de résistance aux protéases pour MB-1Trp. Ainsi, après 60 minutes d'exposition de la protéine MB-1 à la Pronase E, un mélange de protéases bactériennes à activité non-spécifique, aucune protéine MB-1 intacte n'était observable. On pouvait encore observer 60% de protéines MB-1Trp non dégradée pour le même temps d'exposition aux protéases. La mutation s'est donc avérée très efficace (20).

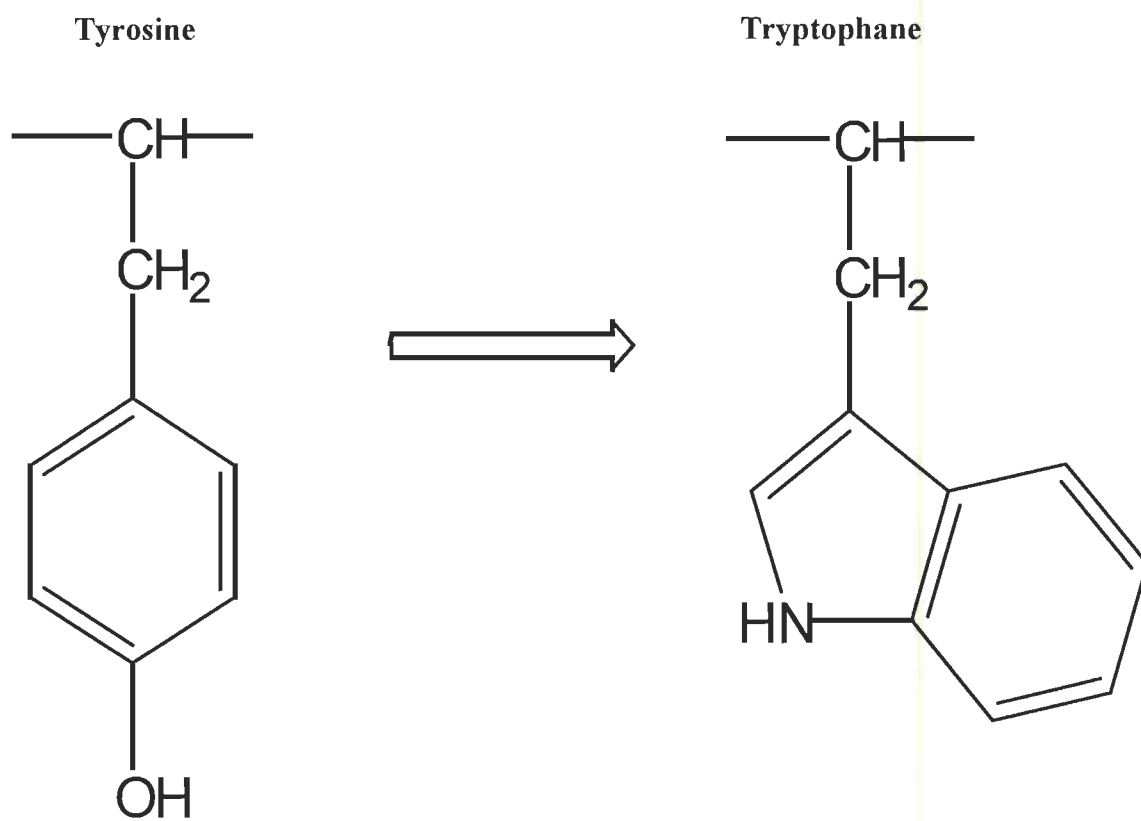


FIGURE 1.12. Structure chimique de la tyrosine et du tryptophane.

1.4 Mon projet de maîtrise : combinaison des deux types de mutations au sein de MB-1

Tel que mentionné, notre laboratoire a effectué deux types de mutations sur la protéine MB-1, soit l'insertion de ponts disulfures et le remplacement de la tyrosine 62 par un tryptophane. Ces deux types de modifications ont eu des effets stabilisants sur la protéine. Il serait donc très intéressant de combiner ces mutations au sein d'une même protéine. Le but de mon projet de maîtrise a donc été de créer deux protéines à pont disulfure comportant chacune un tryptophane en position 62. J'ai donc entrepris la création des protéines MB-1TrpLH et MB-1TrpRH. Pour ce faire, il a fallu remplacer la Tyr par un Trp chez MB-1LH et MB-1RH. Les informations à acquérir lors de ce projet étaient multiples. Nous voulions premièrement savoir s'il était possible de compenser la perte de volume et d'hydrophobicité engendrée lors de la création des mutants à pont S-S en augmentant le volume d'un résidu dans l'environnement du pont. La deuxième information d'importance à obtenir est de connaître si l'effet des deux types de mutations serait additif. Nous voulions donc vérifier si l'insertion du Trp chez MB-1LH et MB-1RH allait amener la même hausse de stabilité chez ces mutants que celle observée pour MB-1. Pour arriver à créer et caractériser les mutants MB-1TrpLH et MB-1TrpRH plusieurs étapes ont été nécessaires. Je décrirai ici les principales étapes de la création et caractérisation des mutants. Le lecteur est invité à lire la section *Experimental Procedures* du Chapitre II pour connaître les détails des manipulations qui ont été effectuées. La première étape est la modification du codon de la tyrosine par celui du tryptophane chez le gène de la protéine MB-1RH et MB-1LH. Lesquels gènes sont déjà insérés dans le plasmide pMAL-c2 et prêt à être exprimés. La modification du codon a été effectuée par mutagenèse dirigée par oligonucléotide et la confirmation de la séquence nucléotidique est effectuée par séquençage automatique. Les protéines sont alors ensuite exprimées et purifiées tel qu'expliqué à la section *Experimental Procedure* du Chapitre II. Une caractérisation biophysique des protéines est ensuite effectuée. Cette caractérisation comprend (a) la détermination des structures secondaires des protéines, (b) la détermination de la formation de leur noyau hydrophobe, (c) la détermination de la

fluidité de ces noyaux ainsi que, (d) la mesure de la résistance thermique et protéolytique des protéines. Les résultats de cette caractérisation sont présentés au Chapitre II qui renferme l'article scientifique intitulé «Engineering nutritious proteins: Improvement of stability in the designer protein MB-1 via introduction of disulfide bridges». Les résultats concernant la caractérisation de MB-1TrpLH et MB-1TrpRH y sont présentés et discutés. Le Chapitre III présente ensuite une discussion plus générale et met en contexte les résultats obtenus par rapport aux autres protéines de la famille MB-1. Il y est aussi discuté des résultats obtenus par d'autres groupes de recherche dans le domaine de l'ingénierie des protéines.

1.5 Présentation de l'article scientifique

Comme je l'ai mentionné précédemment, un article scientifique est présenté au Chapitre II de ce mémoire. Cet article s'intitule « Engineering nutritious proteins: Improvement of stability in the designer protein MB-1 via introduction of disulfide bridges» et il a été soumis au journal *Journal of Agricultural and Food Chemistry* pour publication. La confirmation de la réception du manuscrit par l'éditeur a été reçue le 12 juillet 2001, et l'acceptation après révision mineure a été confirmée en septembre 2001. Le texte de cet article est en langue anglaise de façon à respecter les normes du journal. Une bibliographie est présentée à la fin de l'article et elle ne contient que les références citées dans cet article. Le format de cette bibliographie ne respecte pas les règles établies pour la rédaction d'un mémoire traditionnel mais respecte plutôt le format du journal auquel l'article a été soumis pour publication. La bibliographie relative aux autres chapitres de ce mémoire est présentée en page 79. Le document spécifiant le format à être utilisé pour rédiger un article à être soumis dans ce journal est présenté à l'ANNEXE A. Cette annexe contient aussi une copie des lettres envoyées à qui de droit pour obtenir la permission de rédiger ce mémoire en y incorporant un article scientifique.

CHAPITRE II

RÉSULTATS

Ce chapitre présente les résultats obtenus pour les protéines MB-1RH, MB-1LH, MB-1TrpLH et MB-1TrpRH sous forme d'article scientifique soumis pour publication. Il est à noter que les légendes des figures se retrouvent sur une page séparée des figures. Les figures se rattachant à l'article ne sont donc pas numérotées, elles sont toutefois présentées dans le même ordre que les légendes s'y rattachant et qui sont présentées à la section « Figure Caption » de l'article. À la suite de l'article le tableau récapitulatif 2.1 est présenté. Ce tableau sert à mettre les résultats en contexte. J'y présente donc les résultats obtenus pour la protéine mère MB-1, la protéine MB-1Trp ainsi que les quatre mutants à pont disulfure. Les résultats seront discutés au Chapitre III.

2.1 Article scientifique

Engineering nutritious proteins: Improvement of stability in the designer protein MB-1 via introduction of disulfide bridges.

Alain Doucet, Martin Williams, Mylene Gagnon and Marc Beauregard*

Groupe de Recherche en Énergie et Information Biomoléculaire

Département de chimie-biologie

Université du Québec à Trois-Rivières

C.P. 500, Trois-Rivières, Québec

G9A 5H7 Canada

*Correspondence should be addressed to Marc Beauregard, Professor, Département de chimie-biologie, Université du Québec à Trois-Rivières, C.P. 500, Trois-Rivières, Québec, G9A 5H7 Canada, Telephone: (819)376-5052, Fax: (819)376-5084.

Abstract

Protein design is currently used for the creation of new proteins with desirable traits. In our lab we focus on the synthesis of proteins with high essential amino acid content having potential applications in animal nutrition. One of the limitations we face in this endeavour is achieving stable proteins despite a highly biased amino acid content. We report here the synthesis and characterisation of two disulfide bridged mutants derived from our MB-1 designer protein. Both mutants outperformed their parent protein MB-1 with their bridge formed, as shown by circular dichroism, size-exclusion chromatography, thermal denaturation and proteolytic degradation experiments. When the disulfide bridges were cleaved, the mutants behaviour changed: the mutants significantly unfolded, suggesting that the introduction of Cys residues was deleterious to MB-1 folding. In an attempt to compensate for the mutations used, a Tyr62-Trp mutation was performed, leading to an increase in bulk and hydrophobicity in the core. The Trp containing disulfide bridged mutants did not behave as well as the original MB-1 Trp, suggesting that the position 62 might not be adequate for a compensatory mutation.

Introduction

The expense associated with feed additives used for animal production have led to the study of intracellular production of high quality protein by transgenic crops and other organisms, as a mean of obtaining efficient and less costly sources of essential amino acids (EAA) (1,2,3). Three main approaches for improving protein quality in a given organism are being explored (4,5). The first approach involves transferring a gene coding for a high quality protein from one organism to another that is more suitable for farming practices (heterologous expression). Recently, soybean and sunflower albumins have been chosen for their high methionine contents for the development of transgenic crops by major agrobiotech companies (6,7). In this approach, the amino acid composition of a natural protein is predetermined and may not conform to the desired EAA ratio. The second approach involves modifying the genes of an organism so that more selected essential amino acids are inserted in the proteins. Attempts to changing amino acid content, however, often destabilise the protein and/or prevent it from folding, which may jeopardise its recovery (4,8,9). Recently, some success have been reported with this approach by Chui *et al.* 1999 (10) and Hey *et al.* 2001(11). The third approach, chosen here, involves creating a new protein with a biased composition of selected essential amino acids (12,13,14). Theoretically, this strategy allows for a full control of the amino acid composition of the protein and is, thus, an advantage over the previously mentioned options.

Previous efforts to express synthetic proteins with high EAA contents *in vivo* in bacteria (14,15,16) have not yielded high quantities of protein, which is possibly attributable to the lack of structural stability of the final peptide. In order to improve on the properties of such new proteins, we have tried to identify dominant folding principles in

selected natural proteins and attempted to encode them into the new amino acid sequence. The underlying assumption of this strategy is that by encoding information that confers structure, compactness and stability, a protein normally stable *in vivo* could be designed, regardless of its biased composition. The ideal protein we hope to design would: 1- have a strong biased composition, with methionine, lysine and other essential amino acids being dominant (the limiting amino acids for milk protein production were targeted), 2- be expressed *in vivo* as a soluble, stable, full length protein.

At the time of its design (13), creation of this ideal protein was a challenge yet to be met, since most reported advances in protein design had not involved expressed and soluble *in vivo* full length proteins (17). Nevertheless, the investigation of folding principles in α -helical bundle proteins was relatively advanced (18), and because this fold could harbour most of the residues desired, we selected the bundle as our ideal protein. The basic assumptions for the design of the new protein called Milk Bundle-1 (MB-1) were as follows: 1- that the range of possible residues in each recurring position in the helices of the bundle protein would be determined by the physiochemical properties of the amino acid residues and the topological constraints due to their particular location in the tertiary structure, that 2- the driving force for bundle formation would be a “hydrophobic collapse”, where hydrophobic side chains are energetically favourably buried in the interior of the bundle, and that 3- as a result, the loops would be relatively unimportant. The design process was an iterative one in which criteria such as secondary structure propensities, charged residues distribution and other global factors found determinant for folding (19,20,21,22), were subsequently used to refine the initial design (13).

Characterisation of MB-1 indicated that the design process used resulted in the stable expression of a new, largely helical protein enriched in desired amino acids (60% in M, T, K and L). After a first round of design, the MB-1 protein was found to have a folded core and a low affinity for 8-anilino-1-naphthalenesulfonic acid (ANSA) (23). Its behaviour and expression levels *in vivo* were found to be far superior to that obtained from earlier attempts of high essential amino acid polypeptide design (24).

Investigation of MB-1 properties did, however, reveal some flaws. Its melting temperature was found to be very low (39°C), and protease resistance at physiological temperatures was limited, possibly due to partial unfolding (24). Such results suggests that MB-1 would not accumulate in transgenic organisms, where it would be exposed to proteases. Proteolytic degradation experiments have predicted that MB-1 lifetime in rumen conditions will be shorter than that of several natural proteins (25). Thus, feeding MB-1 to ruminants may also be inefficient.

One way of improving the properties of MB-1 is to increase conformational stability, which often correlate with resistance to proteolytic degradation (26,27,28). In order to improve MB-1 stability, additional design cycle are under study. In the absence of an x-ray resolved structure, extensive core redesign and elaborated fold specifying devices had to be ruled out. In view of its apparent simplicity, we chose the insertion of disulfide bridges. This strategy was not only chosen for its simplicity, but for other reasons as well: 1- an intramolecular disulphide bridge would bring stability to the protein by reducing the entropy of the unfolded state (29,30); 2- the disulphide bridge is a well known stabiliser against proteolysis, probably due to its impact on target access by

proteases (31,32,33); 3- its engineering involves only a small modification of the amino acid composition, which is critical in order for MB-1 to remain nutritionally efficient.

This study also involves MB-1Trp, a mutant that was obtained recently (34). This mutant is derived from MB-1, and was obtained the replacement of a core Tyr by Trp, a larger, less polar amino acid. Several protein properties were positively modified by this mutation, including thermostability and resistance to proteolytic degradation which were largely increased (34). Thus, in this paper, we will attempt to build stabilizing disulfide bridges into MB-1 and MB-1Trp.

Experimental Procedures

Preparation of the new mutants

Substitution of Cys in position 13 and 87 of MB-1 (MB-1LH) and positions 10 and 91 of MB-1 (MB-1RH) was performed using the oligo-directed mutagenesis kit « Altered Sites® II » (Promega). The mutational oligonucleotides (**1** for position 13; **2** for position 87, **3** for position 10; **4** for position 91 shown below with the corresponding MB-1 sequences) were purchased from GibcoBRL/Life Technologies, purified using denaturing polyacrylamide gel electrophoresis (PAGE) and phosphorylated.

MB-1 : 5'-ATG ATG ACC ACC CTG TTT AAA ACT ATG-3'

Oligo 1: L13C : 5'-ATG ATG ACC ACC TGC TTT AAA ACT ATG-3'

MB-1: 5'-ACG GCT ACA ACC ATG AAA AAT CAT CTG-3'

Oligo 2: M87C : 5'-ACG GCT ACA ACC TGC AAA AAT CAT CTG-3'

MB-1: 5'-ATG ACC GAC ATG ATG ACC ACC CTG TTT-3'

Oligo 3: M10C: 5'- ATG ACC GAC ATG TGT ACC ACC CTG TTT-3'

MB-1: 5'-ATG AAA AAT CAT CTG CAG AAC TTG ATG-3'

Oligo 4: L91C: 5'-ATG AAA AAT CAT TGC CAG AAC TTG ATG-3'

MB-1: 5'-ATG GCC ACT ACG TAC TTC AAA ACG-3'

Oligo 5: Y62W: 5'-ATG GCC ACT ACG TGG TTC AAA ACG-3'

The Tyr62-Trp mutation lead to MB-1Trp and was achieved directly in the expression vector pMal-c2. The mutational oligonucleotide **5** was used and treated as described above. The mutations were then confirmed by dideoxynucleotide sequencing using T7 Sequenase kit (Amersham Life Science). The mutated MB-1 genes were cloned back in pCMG20 4-X, the expression vector. All Positive clones were checked again by DNA sequencing after cloning.

Protein expression and purification

All mutated proteins were prepared as described in Gagnon et al. (34). After purification, all proteins were checked for purity by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (protein purity >95%). Unexpectedly, during purification, the mutant MB-1TrpRH formed oligomers of various sizes. For this mutant, an additional purification step was used in order to isolate the monomeric proteins, as described in Doucet and Beauregard 2001(35).

Protein quantification and electrophoresis

Protein concentration was determined by the bicinchoninic acid (BCA) assay (Sigma), using bovine serum albumin as standard. The protein was visualised by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) using 12% polyacrylamide-tricine gels, followed by silver nitrate staining. SDS-PAGE experiments were conducted prior to measurements to confirm protein purity.

Quantification of Cys residues and assessment of bridge formation

Chemical modification of the Cys residues by iodoacetamide and iodoacetate was performed as described in Hollecker and Larcher 1989 (36). Comparison of the electrophoretic mobility of the reduced and non-reduced form of the MB-1 mutant was done by SDS-PAGE (35). The Cys were also quantitated by reacting with *p*-hydroxymercurybenzoate (*p*-HMB) and measuring absorbance of *p*-HMB at 260nm with a UV/vis spectrophotometer (Novaspec, Pharmacia).

Conformational investigation by circular dichroism (CD)

Unless specified otherwise, protein samples were prepared at a concentration of 0.6 mg/mL with a phosphate buffer (128mM NaH₂PO₄, pH 6.8). The reductive agent dithiothreitol (DTT) was added to a final concentration of 10mM where specified. The samples were then degassed and equilibrated 20 minutes at 20°C before measurements. Spectra were measured with a Jasco J-720 spectropolarimeter, which was routinely calibrated with a 0.06% (w/v) ammonium (+)-10-camphorsulfonate solution. For measurements in the far-UV region, a quartz cell with a path length of 0.01 cm was used. Ten scans were accumulated at a scan speed of 20 nm per minutes, with data being collected at every nm from 180 to 260 nm. Sample temperature was maintained at 20°C using a Neslab RTE-111 circulating water bath connected to the water-jacketed quartz cuvette. Spectra were corrected for buffer signal and conversion to $\Delta\epsilon_{MRW}$ (mean residue weight) was performed with the Jasco Standard Analysis software. Secondary structure calculations were performed using the CDsstr program developed by Johnson (37), using default settings.

Thermal denaturation

Samples were prepared as described in the preceding section. In order to measure thermostability, temperature was increased from 10 to 95 °C at a rate of 30 °C per hour using a Neslab RTE-11 controlled by the Jasco spectropolarimeter software. CD spectra were collected at every 5°C, from 200 to 260 nm, at a scan speed of 20 nm/min and CD signal at 222nm were collected at 1°C interval. In order to assess reversibility of thermal denaturation, the protein solutions were cooled down at a rate of 30°C per hour, and spectra were measured at 70, 50 and 20 °C.

Thermal stability was calculated assuming a unimolecular, two-state process as previously described (38). The CD signal at 222nm at various temperature was used as the property (y) indicative of the extent of unfolding. In the folded state, the parameter $y = y_f$ and the fraction of folded protein f_f is equal to 1. When the protein is unfolded, the parameter $y = y_u$, and the fraction of unfolded protein f_u is equal to 1. For intermediate states, y is given by $y_f f_f + y_u f_u$. Thus, by measuring y , we can calculate the fraction of protein unfolded: $f_u = (y_f - y)/(y_f - y_u)$. The equilibrium constant for the unfolding process is $K_u = f_u/(1-f_u)$ and melting temperatures (T_m) are obtained at $K_u = 1$ (38).

8-anilinonaphthalenesulfonic acid (ANSA) fluorescence enhancement

Protein concentration was adjusted to 0.2 mg/mL and equilibrated at room temperature (RT) for 1 hour. Then ANSA was added to a final concentration of 10 μ M and equilibrated 5 minutes prior to measurements. Spectra were recorded using an LS50-B Perkin-Elmer fluorometer with an excitation wavelength of 380 nm. Spectra were

collected from 410 to 550 nm. Correction for buffer signal and for the effect of DTT on ANSA was keyed in when applicable.

Fluorescence quenching measurements

Protein concentration of samples was adjusted to 0.2 mg/mL and equilibrated at RT for 1 hour. Where specified DTT was added to a final concentration of 10mM and equilibrated at RT 15 minutes before adding the quencher citrate (final concentration 0.25 M). Control samples were exposed to 0.75 M NaCl in order to keep the same ionic strength as in the samples exposed to 0.25 M citrate. Mutants containing Tyr were excited at 280nm and fluorescence spectra were recorded from 290nm to 330nm. For the Trp mutants, fluorescence was collected from 310 – 400 nm. Quenching of the Trp mutants was performed with 0.3M potassium iodate, and NaCl was used in controls. Spectra were then corrected for buffer signal before analysis. The effect of DTT on fluorescence was corrected when needed.

Proteolytic degradation measurements

Proteins were dialysed against 2000 volumes of borate-phosphate buffer (pH 6.8) at 4°C overnight. Experiments were then carried out as described by Krisnamoorthy *et al.* (39) with few modifications as described in Morrison *et al.* (25). Reactions were stopped by the addition of 2% SDS buffer followed by heating for 3 minutes at 100°C. Samples were then analysed on SDS-PAGE followed by silver nitrate staining. The amount of protein remaining after incubation with proteases was determined by measuring the optical density of each band using the image analysis system Imaging Research MCID. Cytochrome c (horse heart, Sigma cat no.C-7752) was used as a reference protein for all degradation runs in order to correct for possible variations in crude proteolytic activity.

Size exclusion chromatography (SEC)

Proteins were applied to a Beckman Ultraspherogel SEC-200 column at a concentration of 0.6mg/mL. A constant flow of 0.8mL/min was maintained using a Waters HPLC pump. Pharmacia Low Molecular Weight gel filtration calibration standards were used for column calibration.

Results

Design strategy

The putative modifications to MB-1 structure are illustrated in Figure 1. The design strategy used here focused on two aspects: 1- the restrictive effect of a covalent bond between remote residues on the protein as a whole; and 2- the precise location of Cys which permits disulfide bridge formation. By choosing positions as far apart as possible, one can reduce the entropy gain upon unfolding for most of the protein (29,40). Thus, insertion of a bridge between helices I and IV would enclose a larger part of the polypeptide than a bridge involving other helices. Another consideration for using helix I is that this MB-1 segment of sequence is sensitive to proteolytic degradation (Hefford and Beauregard, unpublished). The restriction of helix I by Cys insertion could help prevent such a phenomenon.

The position of Cys in helices I and IV must allow sulfhydryl groups to be properly aligned in order to minimise strain induced by bridge formation. On the basis of geometric models built for similar proteins (41), it appeared that position "d" of the heptad pattern used for MB-1 design would offer the best geometry for bridge formation. Therefore, L13 and M87 residues were selected for mutation to Cys. Figure 1 depicts the expected location of the bridge in the mutant (hereafter referred to as MB-1LH, assuming it folds as per design). Note that for proper alignment of position "d" in helices I and IV, a left-hand connectivity of the helices had to be assumed (i.e. the bundles are positioned such that when helix I is at the fore front, with its N-terminus pointing down, then helix II is placed to the left of helix I). A second scenario was considered, in which a right-hand connectivity could be specified. Examination of the second model in Figure 1 suggests mutations at positions "a" in helices I and IV, since

positions "d" would be too far apart. By choosing M10 and L91 residues for mutation to Cys, we attempted to generate a mutant (named MB-1RH) that would resemble MB-1LH as much as possible, except for reversing its connectivity.

Disulfide bridges will also be inserted into MB-1Trp. This protein is a derivative of MB-1 where Tyr62 was replaced by Trp. Position 62 in MB-1 was chosen for the emplacement of a spectroscopic probe at the moment of initial design (13). As shown on the model in Figure 2, position 62 is part of the hydrophobic core, and a niche made of 5 Ala was built around it in order to accommodate a larger side chain in this region of the core. The replacement of Tyr by Trp was thought to improve on stability, and indeed, characterisation of MB-1Trp confirmed the strategy. MB-1Trp has a melting temperature of 55°C and is more resistant to protease action than MB-1. Here we are going to use MB-1Trp because of the increase in bulk offered by Trp in the core, in a way to compensate for the loss of volume consequent to the mutations used for bridge insertion.

Initial characterisation of mutant structure

The presence of two Cys in all mutants was confirmed by Cys derivatization with mixtures of iodoacetamide and iodoacetate, and by reacting Cys with *p*-HMB, in agreement with the DNA sequences of the expression vectors (data not shown). Formation of the bridges was confirmed by a comparison of the protein migration after and before treatment with DTT, a reductive agent. Reduction of the protein changed its migration speed, and the impact of reduction appeared to modify mobility of all proteins treated (not shown). Replacement of Tyr by Trp was confirmed by spectrofluorometry.

Secondary structures analysis

Secondary structure analysis of the mutant was performed using circular dichroism (Figure 3). The spectra measured for all mutants with their bridge closed (no DTT) were typical of helical proteins, and calculations of helical amide contents were in the range expected for these proteins (Table 1). MB-1LH, MB-1RH and MB-1TrpLH appeared to have similar secondary structures as their parent molecules MB-1 and MB-1Trp (which contain approximately 50% helical amides (23,34). The protein MB-1 appeared to be tolerant to bridge insertion, regardless of the connectivity imposed by this bridge (either handedness).

Inserting the bridges leading to a right-hand connectivity into MB-1Trp led to a different scenario: the helical content of MB-1TrpRH differed significantly in secondary structure contents, with only 42% helical content (Table 1). The right-hand connectivity in MB-1TrpRH imposes a different conformation, with less helical content.

The mutants were then analysed in the presence of DTT (Figure 4), in order to separately assess the effect of the mutations (insertion of Cys) from the impact of bridge formation. All mutants helical contents dropped when exposed to DTT as indicated by a loss of signal intensity at 190 and 225 nm, with MB-1LH being nearly completely unfolded (Table 1). This suggests that in the absence of the bridge, mutations used here are destabilising. Among the four mutants presented here, MB-1LH had the largest drop in helical content after bridge opening (down to 18%, see Table 1). This means that the mutations M13-Cys and L87-Cys promote unfolding of the protein, and that formation of the bridge compensates for this deleterious effect.

When the same treatment was performed on MB-1TrpLH, the loss of helical content was less severe (down to 34%) (Table 1). The fact that MB-1TrpLH stay

somewhat better folded (34%) than MB-1LH in the presence of DTT may suggest that Trp compensate for the loss of bulk consequent to Cys insertion in the core with a left-hand connectivity. Partial unfolding after DTT treatment was confirmed by fluorescence quenching and ANSA binding for all mutants (not shown).

Reversibility of DTT-induced unfolding was verified by CD measurements. After DTT removal, CD spectra were similar to the native spectra, suggesting that all four proteins refolded after bridge repair (Figure 3).

Quaternary structure

The ability to control protein association and aggregation has been (and still is) an important tumble stone in *de novo* protein design. Thus, we monitored protein association for our mutants under two sets of conditions. First, quaternary organisation was monitored by SEC under benign buffer conditions in a way to observe the impact of both covalent and non-covalent bonds. Then, the proteins were denatured and migrated on an SDS-PAGE gel in order to detect any inter-molecular covalent bond. Results in Table 1 indicate that all mutants, except MB-1TrpRH, were monomeric, regardless of the treatments used here. This confirms that intra-molecular bridges are formed, and suggest that the three proteins fold as planned. At variance, MB-1Trp RH was found to contain 5-10% dimer that resisted SDS treatment, which indicate that these dimers are formed via an inter-molecular bridges. Other oligomerization states have been observed for this protein under specific experimental conditions (35). An increase in surface hydrophobicity was detected using ANSA binding measurements (ratio I_{480} / I_{510} for MB-1TrpRH = 12, compared to a ratio of 7 – 8 for the other three mutants), which indicate that MB-1TrpRH core is more fluid, allowing for inter-molecular bridges.

Effect of the mutations on conformational stability

Conformational stability of both mutants was measured using CD as described in Materials and Methods. Figure 5 shows the four denaturation curves, while calculated T_m values are listed in Table 1. Both MB-1 mutants clearly outperform MB-1 ($T_m = 39^\circ\text{C}$ (24)), by approximately 10°C in thermostability. Approximately 100% of the helical content was recovered after renaturing the protein by cooling to room temperature as described earlier in Materials and Methods, whereas the MB-1 parent molecule could not be refolded under similar experimental conditions (24).

Insertion of bridges into MB-1Trp resulted in a completely different scenario, both mutants being less stable than their ancestor molecule MB-1Trp ($T_m = 55^\circ\text{C}$, Gagnon *et al.* (34)), with MB-1TrpLH being the least stable ($T_m = 42^\circ\text{C}$). The mutants are also less stable than the pair MB-1LH and RH, suggesting that the compensatory mutation Tyr62-Trp is in fact not compatible with the conformations specified by the bridges.

Effect of the mutations on proteolytic degradation

Proteolytic degradation experiments were carried out as described in Figure 6. Both mutants derived from MB-1 outperformed MB-1, with about 40% intact protein left after a 1 hour treatment with proteases (MB-1 cannot be detected after same treatment, see reference 25). MB-1LH is more resistant than MB-1RH in the first 45 min, but the difference decreases after 60 min. The mutants derived from MB-1Trp were less stable than their parent molecule (60% intact MB-1Trp left under same

conditions, see reference 34). MB-1TrpRH behaved like MB-1RH and the other MB-1 mutants, while MB-1TrpLH has the lowest resistance, with 17% proteins left after treatment (Table 1). Once again, the insertion of Trp near the bridge did not lead to any improvement on stability. Degradations were performed in the presence of DTT and higher degradation rates were measured. However, difficulties were encountered due to the inhibitory effect of DTT on the proteolytic activity and the need to prevent re-oxidation of the disulfide bridge, thus strongly limiting the reliability of these data (data not shown).

Discussion

The insertion of disulfide bridges in MB-1 was found to have an important impact on protein behaviour. Both mutant proteins were more resistant and stable than MB-1 (25). Their unfolding (be it after bridge reduction or due to high temperature) was found to be reversible, and their apparent size in solution indicates they are monomeric. Both strategies led to serious behaviour improvement when compared to the MB-1 parent molecule, in accordance to the known impact of disulphides on protein resistance (31,32,33). This phenomenon cannot be ascribed to Cys mutations *per se* (which are destabilising), but is a consequence of specifying connectivity and limiting protein skeleton freedom through disulfide bond formation.

Predicting (and understanding) the impact of Cys mutations separately from the effect of bridge formation have been difficult in the past. As a result, engineering disulfides in natural proteins have sometimes led to loss of stability (29,42,43). Here we attempted to separately monitor the effect of mutation from that of bridge formation. Results obtained in the presence of DTT indicate that Cys mutations are destabilizing for both proteins. A negative impact of Cys insertion in heptad positions "a" and "d" is not unexpected. Cys is a poor helix former, it is smaller than Met or Leu, and its polarity is higher than that of Met or Leu (44,45). Looking back at the design strategy, it appears that the choices made for the Cys mutations may be questioned on several accounts: Cys does not have a property profile (size, hydrophobicity) that perfectly matches the consensus profile for positions 10 and 91, where leucines are the preferred residues; Cys is a poor helix former (46); and finally, the positions chosen are part of the core, where a protein is most sensitive to mutations.

Closing the bridge led to an increase in stability, which suggests that possible strain induced by cross-linking helix I and IV is rather unimportant. Thus, the geometry and the position of the Cys residues appear to be adequate for bridge formation as expected per design.

In MB-1Trp, the core Tyr62 has been replaced by Trp, leading to an increase in hydrophobicity and bulk in the core. Thus, we attempted to compensate the deleterious impact of the Cys mutations by using MB-1Trp for target. The disulfide bridged constructions harboring Trp 62 analysed here were not better behaved than the ones based on MB-1. Further, comparing MB-1TrpRH or MB-1TrpLH to MB-1Trp reveals that inserting the bridges was destabilising. MB-1TrpRH appeared to form concatemers linked by inter-molecular bridges, indicating a high conformational flexibility. MB-1TrpLH had the lowest resistance to proteases, suggesting some undetected flaw in folding. Therefore, the insertion of Trp in position 62 of bridged MB-1 protein is not appropriate, and one might infer that this position is too close to the bridges, promoting steric hindrance. Also, loss of backbone flexibility by insertion of the bridge can lead to a misalignment of the Trp residue in its “niche” formed by the 5 surrounding alanines. More definitive answers to this question will be available when we get access to more refined structural data.

Since we did not explicitly specify connectivity in MB-1 design, it would have been tempting to elucidate this topological feature based on a comparison of MB-1LH and MB-1RH (the two connectivities are shown in Figure 1). Unfortunately, with their bridge closed, the proteins behave in a similar manner. The slight differences which would indicate that MB-1RH is better folded may be readily explained by the impact of Cys mutations on MB-1LH. The opposing effect of the bridge and the mutations compromises any attempt at elucidating connectivity. Thus, our results do not clearly indicate any intrinsic preference for a right- or left-hand connectivity in the MB-1 parent molecule. The original MB-1 design was based on a statistical survey of helices in four

Literature Cited

1. Dyer, J.M.; Nelson, J.W.; Murai, N. Strategies for selecting mutation sites for methionine enhancement in the bean seed storage protein phaseolin. *J. Prot. Chem.* **1993**, *12*, 545-560.
2. Ealing, P.M.; Hancock, K.R.; White, D.W. Expression of the pea albumin 1 gene in transgenic white clover and tobacco. *Transgenic Res.* **1994**, *3*, 344-354.
3. Williamson, M. Can the risks from transgenic crop plants be estimated? *TIBTECH.* **1996**, *14*, 449-450.
4. DeLumen, B.O.; Krenz, D.C.; Revilleza, J. Molecular strategies to improve the protein quality of legumes. *Food Technol.* **1997**, *51*, 67-70.
5. Hefford, M.A. Engineering nutritious proteins. *Biotechnol. Genet. Eng.* **1997**, *14*, 191-210.
6. Beach, L.; Tarczynski, M.C. Methods of increasing accumulation of essential amino acids in seeds. **2000**, U.S.A patent No. 6,127,600.

7. Jung, R.; Hastings, C.; Coughlan, S.; Hu, D. Alteration of amino acids compositions in seeds. **1998**, U.S.A. patent No. 5,850,016.
8. Hoffman, L.M.; Donaldson, D.D.; Herman, E.M. A modified storage protein is synthesized, processed, and degraded in the seed of transgenic plants. *Plant Mol. Biol.* **1988**, *11*, 717-729.
9. Lawrence, M.C.; Suzuki, E.; Varghese, J.N.; Davis, P.C.; Van Donkelaar, A.; Tulloch, P.A.; Colman, P.M. The three-dimensional structure of the seed storage protein phaseolin at 3 Å resolution. *EMBO J.* **1990**, *9*, 9-15.
10. Chui, C-F.C.; Falco, S.C.; Rice, J. A.; Knowlton, S. High sulfur seed protein gene and method for increasing the sulfur amino acid content of plants. **1999**, U.S.A patent No. 5,939,599.
11. Hey, T.D.; Merlo, A.O.; Walsh, T. A. Nucleotide sequences of genes encoding sink protein and uses thereof for improving the nutritional quality of feeds. **2001**, U.S.A patent No. 6,169,232
12. Falco, S.C.; Keeler, S.J.; Rice, J.A. Synthetic storage proteins with defined structure containing programmable levels of essential amino acids for improvement of the nutritional value of plants. **1996**, U.S.A. patent No. 5,559,223.

13. Beauregard, M.; Dupont, C.; Teather, R.M.; Hefford, M.A. Design, expression, and initial characterization of MB1, a de novo protein enriched in essential amino acids. *Biotechnology* **1995**, *13*, 974-981.
14. Jaynes, J.; Langridge, P.; Anderson, K.; Bond, C.; Sands, D.; Newman, C.; Newman, R. Construction and expression of sythetic DNA fragments coding for polypeptides with elevated levels of essential amino acids. *Appl. Microbiol. Biotechnol.* **1985**, *21*, 200-205
15. Doel, M.T.; Eaton, M.; Cook, E.A.; Lewis, H.; Patel, T.; Carey, N.H. The expression in E. coli of sythetic repeating polymeric genes coding for poly(L-aspartyl-L-phenylalanine). *Nucleic Acids Res.* **1980**, *8*, 4575-4592.
16. Beauregard, M.; Hefford, M.A.; Teather, R.M. Detection of beta-galactosidase alpha-peptide fusions using a commercially available monoclonal antibody. *Biotechniques.* **1994**, *16*, 831-838.
17. Kohn, W.D.; Hodges, R.S. De novo design of alpha-helical coiled coils and bundles: models for the development of protein-design principles. *TIBTECH.* **1998**, *16*, 379-389.
18. Betz, S.F.; Raleigh, D.P.; Degrado, W.F. De novo protein design - From molten globules to native-like states. *Curr. Opinion Struct. Biol.* **1993**, *3*, 601-610.

19. Cohen, C.; Parry, D.A. Alpha-helical coiled coils and bundles: how to design an alpha-helical protein. *Proteins*. **1990**, *7*, 1-15.
20. Goraj, K.; Renard, A.; Martial, J.A. Synthesis, purification and initial structural characterization of octarellin, a de novo polypeptide modelled on the α/β -barrel proteins. *Prot. Eng.* **1990**, *3*, 256-266.
21. Beauregard, M.; Goraj, K.; Heremans, D.; Martial, J.A. Spectroscopic investigation of structure in octarellin (a de novo protein designed to adopt the α/β -barrel packing). *Prot. Eng.* **1991**, *4*, 745-749.
22. Hill, C.P.; Handeston, D.H.; Wesson, L.; DeGrado, W.F.; Eisenberg, D. Crystal structure of $\alpha 1$: Implications for protein design. *Science*. **1990**, *1249*, 543-546.
23. Hefford, M.A.; Dupont, C.; MacCallum, J.; Parker, M.H.; Beauregard, M. Characterization of MB-1. A dimeric helical protein with a compact core. *Eur. J. Biochem.* **1999**, *262*, 467-474.
24. MacCallum, J.D.; Hefford, M.A.; Omar, S.; Beauregard, M. Prediction of folding stability and degradability of the de novo designed protein MB-1 in cow rumen. *Appl. Biochem. Biotechnol.* **1997**, *66*, 83-93.

25. Morrison, J.J.; McNally, J.D.; Navidzadeh, A.; Beauregard, M. Development of an optimized feeding technology for dairy cows: improvement in resistance to ruminal proteases in the de novo-designed protein MB-1. *Appl. Biochem. Biotechnol.* **2000**, *87*, 247-264.
26. Huang, X.L.; Catignani, G.L.; Swaisgood, H.E. Relative structural stabilities of beta-lactoglobulins A and B as determined by proteolytic susceptibility and differential scanning calorimetry. *J. Agric. Food Chem.* **1994**, *42*, 1276-1280.
27. Parsell, D.A.; Sauer, R.T. The structural stability of a protein is an important determinant of its proteolytic susceptibility in *Escherichia coli*. *J. Biol. Chem.* **1989**, *264*, 7590-7595.
28. Liao, H.H. Thermostable mutants of kanamycin nucleotidyltransferase are also more stable to proteinase-K, urea, detergents, and water-miscible organic solvents. *Enzyme Microb. Technol.* **1993**, *115*, 286-292.
29. Betz, S.F. Disulfide bonds and the stability of globular proteins. *Protein Sci.* **1993**, *2*, 1551-1558.
30. Wedemeyer, W.J.; Welker, E.; Narayan, M.; Scheraga, H.A. Disulfide bonds and protein folding. *Biochemistry.* **2000**, *39*, 7032.

31. Mahadevan, S.; Erfle, J.D.; Sauer, F.D. Degradation of soluble and insoluble proteins by *Bacteroides amylophilus* protease and by rumen microorganisms. *J. Anim. Sci.* **1980**, *50*, 723-728.
32. McNabb, W.C.; Spencer, D.; Higgins, T.J. In-vitro rates of rumen proteolysis of ribulose-1,5-biphosphate carboxylase (Rubisco) from lucerne leaves, and ovalbumine, vicilin and sunflower albumin 8 storage proteins. *J. Sci. Food Agric.* **1994**, *64*, 53-59.
33. Spencer, D.; Higgins, T.J.; Freer, M.; Dove, H.; Coombe, J.B. Monitoring the fate of dietary proteins in rumen fluid using gel electrophoresis. *Br. J. Nutr.* **1988**, *60*, 241-247.
34. Gagnon, M.C.; Williams, M.; Doucet, A.; Beauregard, M. Replacement of tyr62 by trp in the designer protein milk bundle-1 results in significant improvement of conformational stability. *FEBS Lett.* **2000**, *484*, 144-8.
35. Doucet, A.; Beauregard M. A sodium dodecyl sulfate-polyacrylamide gel electrophoresis protein ladder made of disulfide-bridged proteins. *Anal. Biochem.* **2001**, *292*, 296-297.
36. Hollecker, M.; Larcher, D. Conformational forces affecting the folding pathways of dendrotoxins I and K from black mamba venom. *Eur. J. Biochem.* **1989**, *179*, 87-94.

37. Johnson, W.C. Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins*. **1999**, *35*, 307-312.
38. Pace, C.N.; Shirley, B.A.; Thomson, J.A. Measuring the conformational stability of a protein. In *Protein Structure a Practical Approach*; Creighton, T.E.; Information Press Ltd, Oxford, England. 1989, 311-330.
39. Krishnamoorthy, U.; Sniffen, C.; Stern, M.; Van Soest, P. Evaluation of a mathematical model of rumen digestion and an in vitro simulation of rumen proteolysis to estimate the rumen-undegraded nitrogen content of feedstuffs. *Brit. J. Nutri.* **1983**, *50*, 555-568.
40. Klink, T.A.; Woycechowsky, K.J.; Taylor, K.M.; Raines, R.T. Contribution of disulfide bonds to the conformational stability and catalytic activity of ribonuclease A. *Eur. J. Biochem.* **2000**, *267*, 566-72.
41. Zhou, N.E.; Kay, C.M.; Hodges, R.S. Disulfide bond contribution to protein stability - positional effects of substitution in the hydrophobic core of the 2-stranded alpha-Helical coiled-coil. *Biochemistry*. **1993**, *32*, 3178-3187.
42. Hinck, A.P.; Truckses, D.M.; Markley, J.L. Engineered disulfide bonds in staphylococcal nuclease: effects on the stability and conformation of the folded protein. *Biochemistry*. **1996**, *35*, 10328-10338.

43. Vogl, T.; Brengelmann, R.; Hinz, H.J.; Scharf, M.; Lotzbeyer, M.; Engels, J.W. Mechanism of protein stabilization by disulfide bridges: calorimetric unfolding studies on disulfide-deficient mutants of the alpha-amylase inhibitor tendamistat. *J. Mol. Biol.* **1995**, *254*, 481-496.
44. Koehl, P.; Levitt, M. Structure-based conformational preferences of amino acids. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 12524-12529.
45. Saunders, A.J.; Young, G.B.; Pielak, G.J. Polarity of disulfide bonds. *Protein Sci.* **1993**, *2*, 1183-1184.
46. Blaber, M.; Lindstrom, J.D.; Gassner, N.; Xu, J.; Dirk, W.H.; Matthews, B.W. Energetic cost and structural consequences of burying a hydroxyl group within the core of a protein determined from Ala->Ser and Val->Thr substitutions in T4 lysozyme. *Biochemistry.* **1993**, *32*, 11363-11373

Figure Captions

Figure 1. Schematic representation of MB-1LH and MB-1RH. The two ribbon models show the disulfide bridge (tube connecting helices) to be formed between helices I and IV. Top left: ribbon model of an α -helical bundle with a left-hand connectivity. Bottom left: end-on view showing the canonical heptad pattern used to design MB-1. Positions 'd' are chosen for mutations to Cys for MB-1LH. Pictograms on the right side of the figure apply to MB-1RH (with a right-hand connectivity) where positions 'a' are selected for the mutations.

Figure 2. MB-1's hydrophobic core as predicted by design. The intended structure is a four-helix bundle, with residue polarity distributed along the canonical heptad pattern. The pictogram shows only the interior of the bundle after 'opening' and 'flattening', each box representing the non-polar face of a helix. Note that the large Tyr in position 62 is surrounded by five small alanine residues. This Tyr62 is changed to Trp in MB-1Trp.

Figure 3. CD spectra of the four mutants. CD signal is expressed as change in extinction coefficient per residue. Open squares: MB-1RH; open circles: MB-1LH; closed squares: MB-1RHTrp; closed circles: MB-1LHTrp. Note that after a cycle of reduction-oxidation or after renaturation, the CD spectra measured superimposed on the ones shown here (see Materials and Methods for details).

Figure 4. CD spectra measured in the presence of DTT. Symbols are like those in Fig. 3.

Figure 5. Thermal denaturation curves. The proteins are folded (Fraction unfolded = 0) at low temperature, and unfolded at high temperature (Fraction unfolded = 1). For symbols see Figure 1.

Figure 6. Proteolytic degradation of the proteins. Percentages are based on total protein (MB-1 or mutant) detected prior to incubation. Symbols are like those in previous figures, and results obtained for cytochrome c are shown by '×'.

Table 1. Comparison of physical properties measured for the four mutants.

Mutant	α -helix ^a +/- 3 %	Monomer population SEC ^b +/- 5 %	Monomer population SDS-PAGE ^c +/- 5 %	T_m ^d +/- 1 ^o C	Protein undegraded ^e +/- 8 %
RH	50 (23) ^f	100	100	50	35
LH	48 (18)	100	100	48	43
RHTrp	42 (25)	90	95	49	40
LHTrp	55 (34)	100	100	42	17

a : Helix percentage calculated with CD spectra.

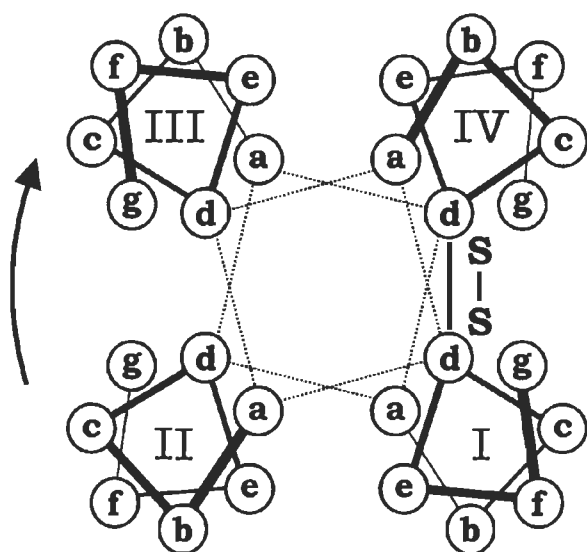
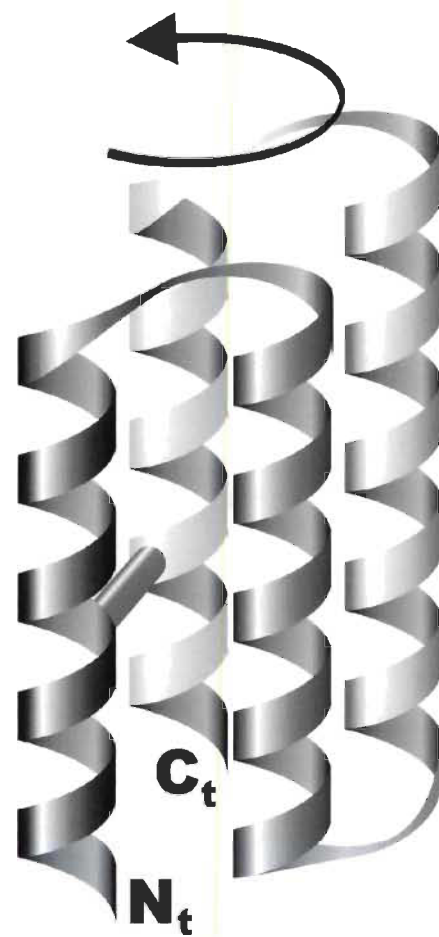
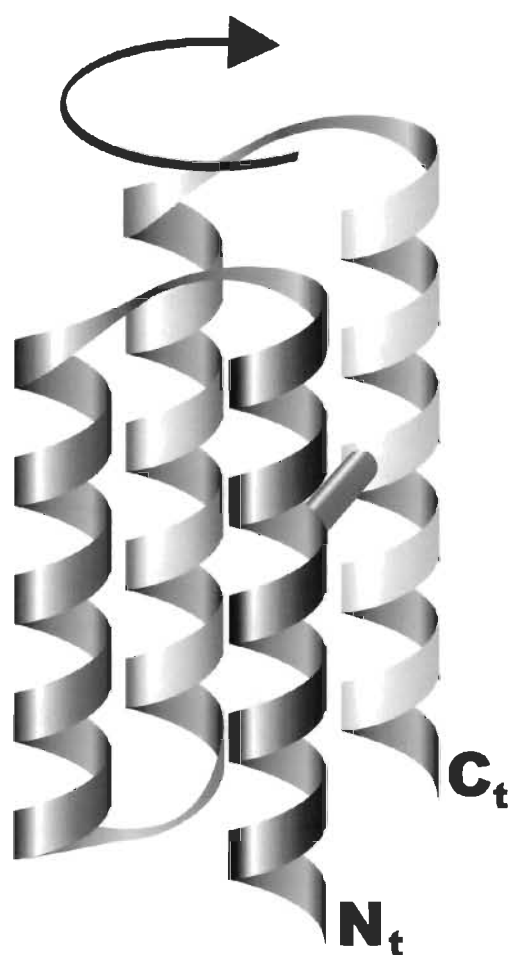
b : Oligomerization population estimated from exclusion chromatography .

c : Oligomerization population estimated from SDS-PAGE gel analysis

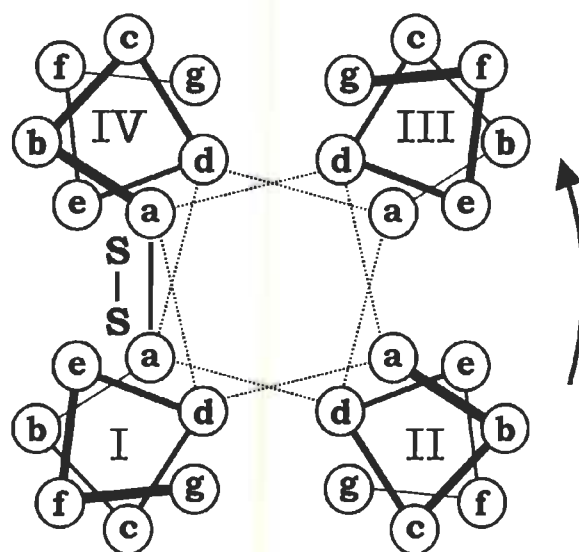
d : Melting temperature derived from thermal denaturation.

e : Percentage of protein undegraded after a 60 minutes of incubation with pronase E.

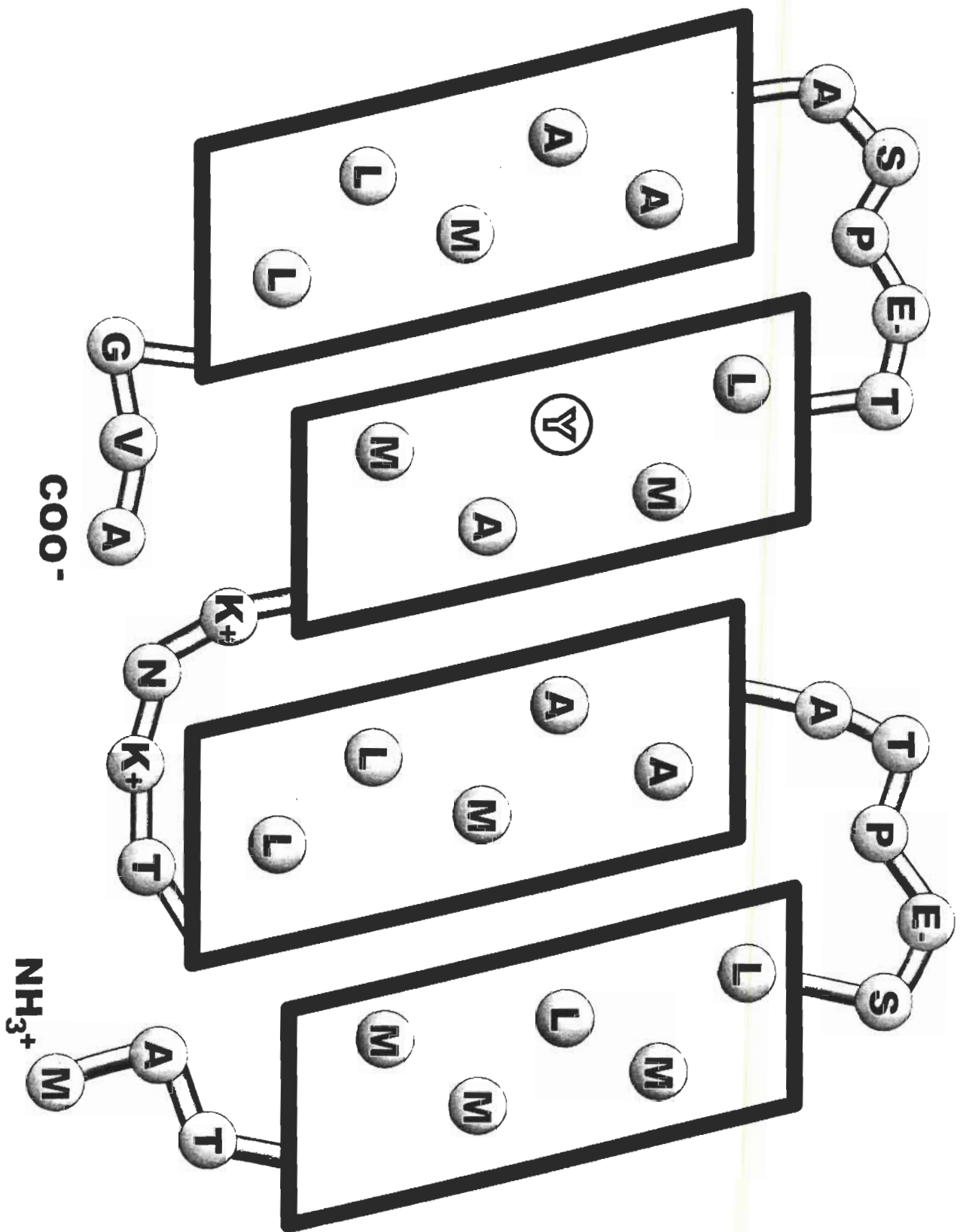
f : Numbers in parenthesis are for proteins with opened disulphide bridge (treated with DTT).

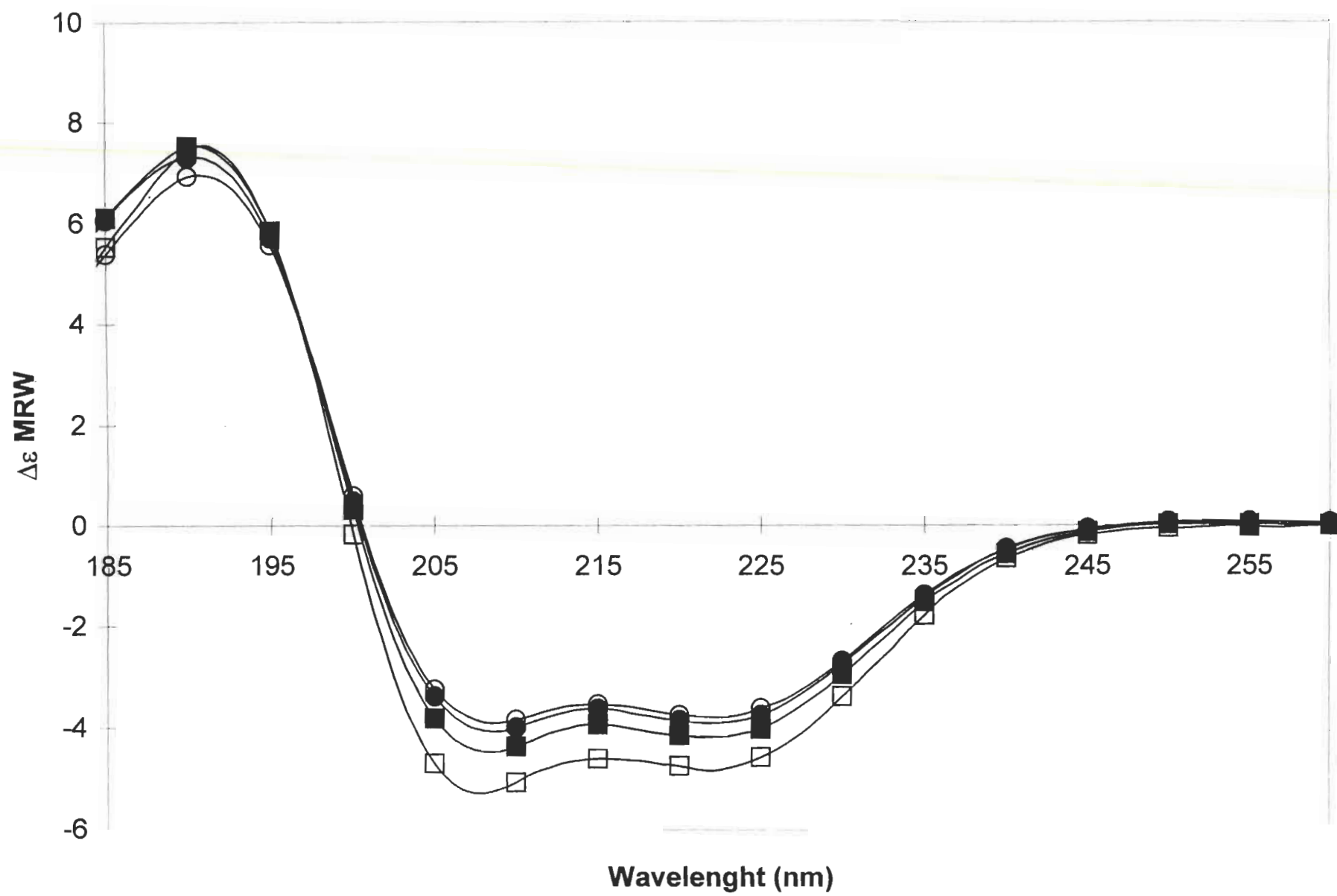


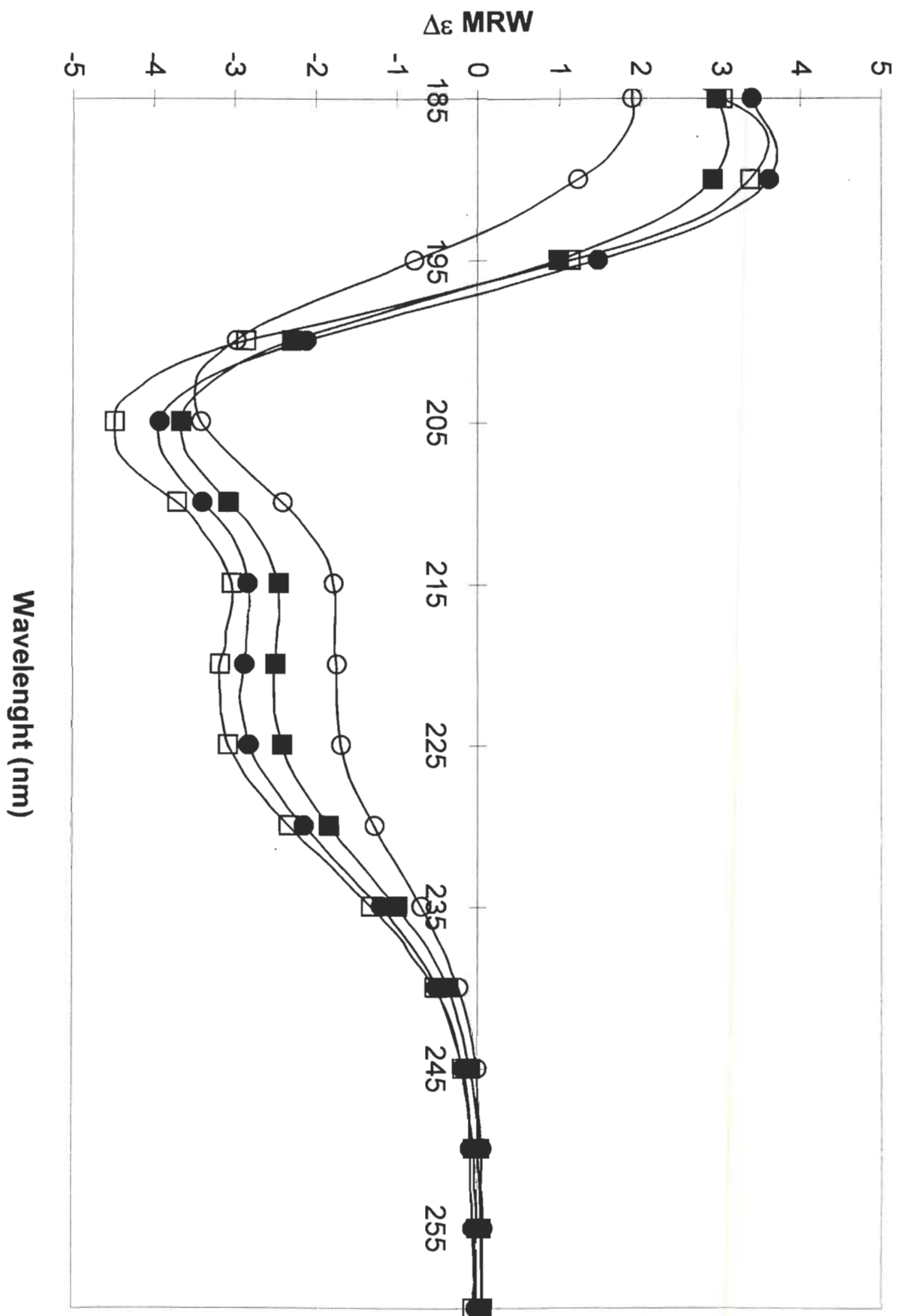
Left-Handed

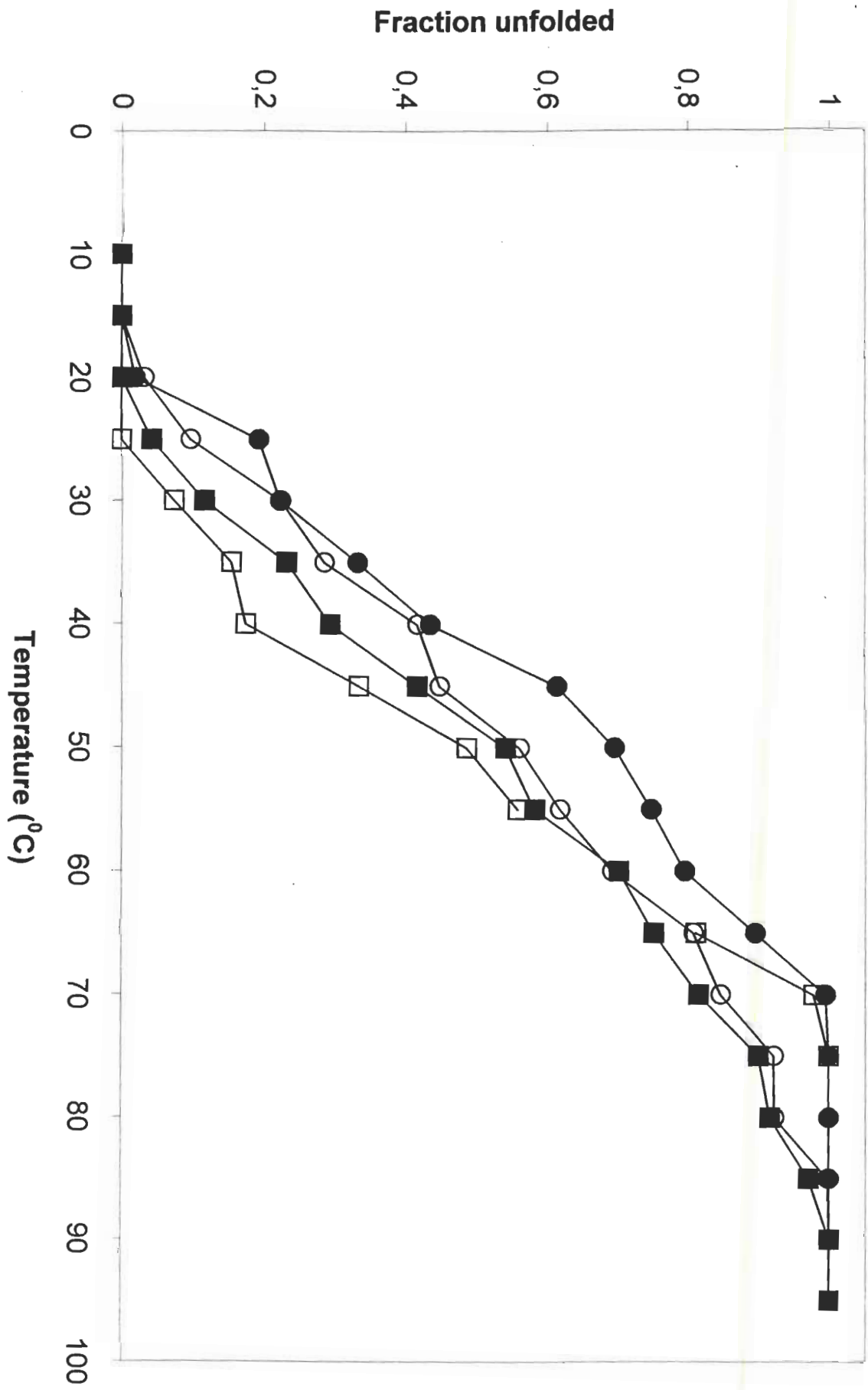


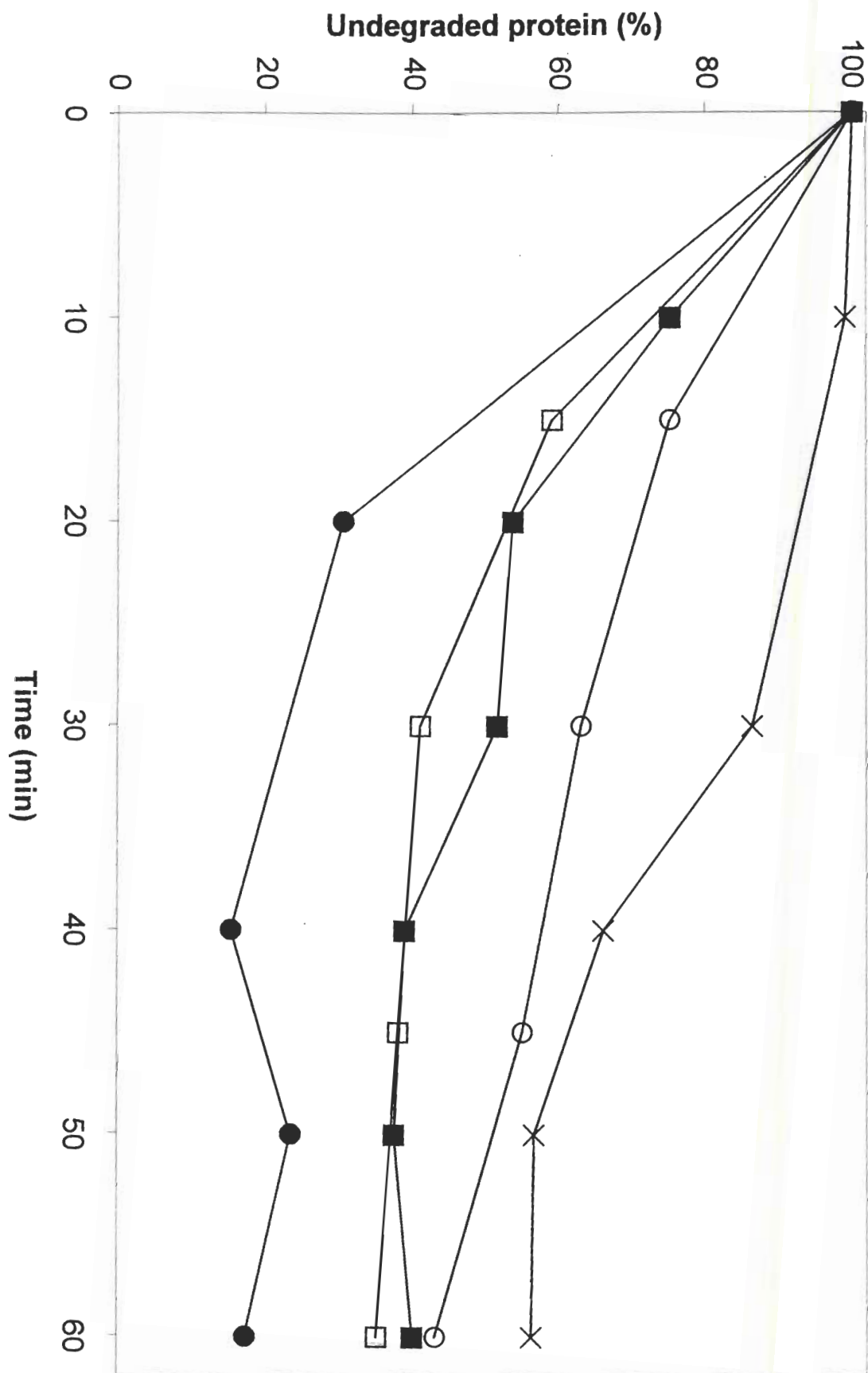
Right-Handed











2.2 Tableau Comparatif

TABLEAU 2.1

Tableau comparatif des propriétés physiques de MB-1 et de différents mutants.

	% hélice α sans DTT	% hélice α avec DTT	Fluidité du noyau (ratio I_{480nm})	λ max du Trp (nm)	T_m ($^{\circ}C$)	Résistance aux protéases (% non dégradée après 60 min)
MB-1	50	N.A.	3	N.A.	39	0
MB-1Trp	50	N.A.	3	333	55	60
MB-1LH	48	18	8,6	N.A.	48	43
MB-1TrpLH	55	34	7,0	345	42	17
MB-1RH	50	23	5,6	N.A.	50	35
MB-1TrpRH	42	25	11,9	345	49	40

CHAPITRE III

DISCUSSION

3.1 Remise en contexte de notre travail

Le but de cette discussion est de mettre en relief les résultats obtenus pour les protéines MB-1TrpLH et MB-1TrpRH que j'ai personnellement produites et caractérisées. Pour analyser clairement les résultats obtenus il est d'intérêt de tout d'abord connaître à fond le contexte de travail de notre laboratoire.

3.1.1 Type de protéine produite et méthode de production

Il faut premièrement savoir que nous travaillons sur une protéine pleine longueur avec un fort biais au niveau de sa séquence en A.A.. La plupart des groupes travaillant dans le domaine de l'ingénierie des protéines travaillent avec des peptides ne comportant qu'une ou deux hélices. La séquence de ces hélices est choisie de façon à ce qu'elles s'associent en dimère, trimère, etc. Ces peptides sont utilisés pour connaître l'effet de la séquence sur la structure de l'oligomère et n'ont habituellement aucune application d'ordre pratique. Aucun biais au niveau de la séquence n'est alors imposé lors du choix de la séquence primaire, ceci facilite de beaucoup le travail. De plus, ces hélices présentent souvent une quantité impressionnante de ponts salins (21). Cette caractéristique permet aux chercheurs de spécifier l'orientation relative des hélices ainsi que l'état d'oligomérisation qu'ils veulent que le peptide adopte. Ces stratégies s'avèrent

efficaces et des nœuds de quatre hélices très stables ont été produits de cette façon. Toutefois, la structure de ces polypeptides ne concorde pas avec la structure des protéines naturelles du fait qu'à ma connaissance aucune protéine naturelle ne présente autant de ponts salins. Les protéines MB-1TrpLH et TrpRH se rapprochent plus des protéines naturelles que ces peptides synthétiques.

La méthode de production des peptides utilisés par d'autres groupes diffère aussi beaucoup de notre mode de production. Les peptides sont souvent obtenus par synthèse chimique. Ils ne sont en contact avec aucun organisme vivant. La synthèse chimique n'est cependant pas pratique pour une protéine de cent acides aminés comme les nôtres car la probabilité qu'une erreur se produise lors de la synthèse augmente avec la longueur de la chaîne à synthétiser. Notre protéine est donc synthétisée par un organisme vivant, soit par la bactérie *E. coli*. Ce type de synthèse implique plusieurs facteurs : il faut que la séquence de la protéine soit encodée dans un gène qui sera inséré dans la bactérie pour qu'il y ait production de la protéine. Ce facteur implique qu'il faut surmonter les problèmes d'instabilité génétique reliée à l'insertion d'un nouveau plasmide au sein d'une bactérie. Il faut de plus changer la séquence du gène pour apporter une modification au niveau de la séquence protéique, ce qui prend un certain temps. Tous ces problèmes sont évités lorsque le peptide est synthétisé chimiquement. Les protéines qui sont produites à l'aide d'organismes vivants sont de plus exposées aux protéases intracellulaires de ces organismes. Il est connu qu'une protéine ne possédant pas les caractéristiques d'une protéine native, et c'est le cas pour plusieurs protéines dessinées *de novo*, résiste mal à la dégradation protéolytique intracellulaire chez *E. coli*. (22). Ceci amène donc un rendement de production très bas car les protéines sont dégradées au sein de la cellule et en partie irrécupérables. Ce problème aussi est évité lors de la synthèse chimique. La purification d'un peptide issu d'une synthèse chimique est très facile puisque que l'on doit retirer que les réactifs et produits qui sont reliés à la synthèse. Ces molécules sont généralement de petite taille comparativement à la taille d'un peptide et peuvent être retirées par simple dialyse ou chromatographie d'exclusion de taille. Par contre, une protéine produite et emmagasinée dans le cytoplasme de *E. coli*

doit être séparée de milliers d'autres protéines bactériennes de taille semblable. La purification comprend donc plusieurs étapes et une perte de matériel doit être assumée pour chacune de ces étapes. Ce fait contribue à diminuer le rendement de récupération en plus d'allonger le temps nécessaire à l'obtention d'une protéine pure. Mon projet de recherche impliquait la production et la purification de deux protéines dans des conditions beaucoup plus exigeantes que ce qui prévaut au sein de d'autres laboratoires dans le domaine de l'ingénierie des protéines. Toutefois, ces conditions sont très similaires aux conditions de production d'une protéine naturelle et notre démarche est donc beaucoup plus près de la réalité. Il faut aussi garder en tête qu'ultimement nos protéines pourraient être produites par des plantes modifiées génétiquement, alors si ces protéines ne peuvent être accumulées de façon stable dans des bactéries il est peu probable qu'elles puissent être accumulées chez une plante. On peut aussi associer le taux de production et la capacité d'entreposer nos protéines à leur stabilité structurale. Ainsi, la protéine MB-1 devait être utilisée aussitôt qu'elle était purifiée car on ne retrouvait aucune trace de protéine intacte après une semaine de conservation de la protéine à 4°C. La dégradation par des traces de protéases bactériennes dans la solution de protéines est donc très rapide pour MB-1. En ce qui concerne les quatre mutants à pont disulfure et le mutant MB-1Trp, on les retrouve intactes après plus d'un mois de conservation dans les mêmes conditions. Ce résultat confirme donc une augmentation de la stabilité des mutants par rapport à MB-1. Ce type d'information n'aurait pas pu être obtenu si nous avions utilisé la méthode de synthèse chimique.

3.2 Élément original de mon projet de recherche

Comme je l'ai mentionné précédemment, la majorité des groupes de recherche en ingénierie des protéines travaillent avec des peptides qui s'associent pour former différents oligomères. Ces oligomères sont composés d'hélices similaires ou non qui ne sont pas reliées entre elles par des liens covalents. On ne peut donc pas déterminer l'ordre exact dans laquelle les hélices sont disposées les unes par rapport aux autres.

Pourtant, ce type de détail est très bien spécifié dans la séquence primaire des protéines naturelles puisque la majorité d'entre elles adoptent une conformation unique. Il est donc important de spécifier la topologie d'une protéine *de novo* pleine longueur si l'on veut que cette protéine ait les qualités d'une protéine native. Nous avons donc spécifié la topologie main droite et main gauche chez la protéine MB-1 à l'aide de pont disulfure. Ce type de contrôle fin de la topologie d'une protéine n'a pas, à ma connaissance, encore été rapporté par d'autre groupe de recherche. Ceci ajoute donc un caractère très original à mon projet de recherche.

3.3 Discussion des résultats obtenus

L'article scientifique présente les résultats obtenus lors de la caractérisation des protéines MB-1LH, MB-1RH, MB-1TrpLH et MB-1TrpRH. Les résultats obtenus font l'objet d'une discussion à la section *Discussion* du Chapitre II. Je consacrerai donc le présent chapitre à une analyse plus détaillée des résultats obtenus pour mes deux mutants et comparerai ces résultats à ceux obtenus pour les autres protéines de la famille MB-1. Pour comparer mes résultats nous pouvons adopter trois points de vue différents : Les protéines MB-1TrpLH et TrpRH peuvent être considérées comme ayant été produites par (a) le remplacement de la Tyrosine 62 par un Trp et de l'insertion de ponts disulfures chez MB-1, (b) l'insertion de ponts disulfures au sein de MB-1Trp, (c) le remplacement de la Tyr62 par un Trp chez les protéines MB-1LH et MB-1RH. Ces trois points de vue vont être abordés lors de l'analyse des résultats expérimentaux.

3.3.1 Les mutants TrpLH et TrpRH vs. la protéine MB-1

La première ligne du tableau 2.1 en page 68 montre les résultats obtenus pour la protéine MB-1. On constate que cette protéine a un fort taux de structure hélicale et un

noyau hydrophobe bien formé et peu fluide. Toutefois, cette protéine est très sensible à la dénaturation thermique et à la dégradation protéolytique. Toutes les autres protéines présentées dans ce tableau surpassent la protéine MB-1 en terme de stabilité thermique et protéolytique. Si l'on compare le mutant MB-1TrpLH à MB-1 on constate que ce mutant a une quantité de structures hélicales légèrement supérieur à celui retrouvé chez MB-1. On remarque aussi que ce gain de structure s'accompagne d'une augmentation de la fluidité du noyau hydrophobe par un facteur supérieur à deux chez la protéine. C'est donc dire que les structures formées ont tendance à changer plus rapidement de conformation dans le temps. La température de dénaturation est de 4°C plus élevée pour TrpLH comparativement à MB-1 et ce mutant résiste mieux à la dégradation protéolytique.

En ce qui concerne TrpRH, les résultats sont très différents. TrpRH accuse une perte de structures secondaires comparativement à MB-1 et son cœur hydrophobe est quatre fois plus fluide que celui de MB-1. Malgré cette perte de structure et cette augmentation de fluidité de la protéine, celle-ci a un T_m de 10°C supérieur à MB-1 et résiste aussi beaucoup mieux à la protéolyse. On constate donc que les mutants MB-1TrpLH et TrpRH sont beaucoup mieux adaptés pour être utilisés dans la production animale comparativement à MB-1. Ces protéines devront toutefois être produites dans un organisme où un compartiment cellulaire particulier favorise la formation des ponts disulfures.

3.3.2 Insertion de ponts disulfures chez MB-1Trp

La situation est tout autre lorsque l'on compare TrpLH et TrpRH à la protéine MB-1Trp. Les deux mutants ont un cœur hydrophobe beaucoup plus fluide que celui de MB-1Trp et leur tryptophane respectif est exposé au solvant alors que celui de MB-1Trp est caché dans le cœur hydrophobe de la protéine comme il devrait l'être. MB-1TrpLH

et MB-1TrpRH résiste aussi moins bien à la chaleur et à la dégradation protéolytique que MB-1Trp. L'insertion d'un tryptophane n'a donc pas donné les résultats escomptés. Comme l'addition de pont disulfure à la protéine MB-1 a eu un effet stabilisant (lorsque les ponts sont fermés) l'addition de tels ponts chez MB-1Trp aurait dû amener une augmentation de la stabilité de MB-1Trp, ce qui n'est pas le cas. On peut avancer une explication à ce phénomène.

Ainsi, l'insertion d'une tyrosine dans le cœur hydrophobe de MB-1 a nécessité la formation d'une niche de grand volume pour accommoder ce gros résidu. Lors de la conception de MB-1, des résidus de petite taille, soit des alanines, ont été choisis pour se retrouver dans l'environnement de la tyrosine laissant ainsi beaucoup d'espace pour l'insertion de ce résidu dans le cœur hydrophobe. Le tryptophane des mutants Trp devrait donc aller s'insérer dans cette niche à la place de la tyrosine. Toutefois, il faut un bon alignement des hélices les unes par rapport aux autres pour que le résidu puisse s'insérer convenablement dans sa niche. La formation du pont disulfure crée beaucoup de restriction en dictant la position de l'hélice I par rapport à l'hélice IV. Cet apport de rigidité chez la protéine peut empêcher celle-ci d'adopter un alignement des hélices permettant au tryptophane de s'insérer convenablement dans la niche créée par les alanines. Le tryptophane se retrouverait alors dans un endroit où le cœur hydrophobe de la protéine ne permet pas son insertion sans briser l'intégrité du cœur hydrophobe. Le tryptophane se voit donc contraint d'être exclu du noyau hydrophobe et se retrouve au moins partiellement exposé au solvant. Le Trp ne peut être inséré correctement dans le noyau mais il est énergétiquement défavorable d'exposer ce groupement hydrophobe à la surface de la protéine. Cette protéine doit chercher constamment un compromis conformationnel pour accommoder le Trp. Ceci pourrait expliquer le comportement dynamique de la protéine. Ceci ne reste cependant qu'au stade hypothétique puisque sans l'obtention de la structure atomique de la protéine il nous est impossible d'affirmer avec certitude à quel endroit se trouve le tryptophane chez les protéines MB-1TrpLH et TrpRH. De plus, l'obtention du degré d'exposition du tryptophane nous amènerait une information supplémentaire. On saurait alors si le Trp est exposé totalement ou

partiellement au solvant. Cette information pourrait être obtenue en mesurant le taux de quenching de fluorescence en présence de différentes concentrations d'agent quancheur puis en effectuant une analyse de type Stern-Volmer. Cette analyse n'a pas été effectuée dans le cadre de ce travail et mériterait une attention particulière lors d'analyses futures.

3.3.3 Ajout d'un tryptophane à la protéine MB-1LH et MB-1RH

Dans un troisième temps nous pouvons observer l'effet de l'insertion d'un tryptophane au sein des protéines MB-1LH et MB-1RH en les comparant aux protéines MB-1TrpLH et TrpRH. Plusieurs aspects ont déjà été abordés au Chapitre II et je tenterai d'éclaircir ceux-ci.

Comparons tout d'abord les deux mutants de connectivité main gauche. Nous notons premièrement une augmentation du taux de structures secondaires hélicales lors de l'insertion du Trp au sein de la protéine MB-1LH et ce, que le pont disulfure soit formé ou non. Cet effet d'augmentation de la stabilité est dû à une diminution de l'effet déstabilisant des cystéines présentes dans le cœur hydrophobe de la protéine. Le Trp diminue donc la perte de structures lors de l'ouverture du pont S-S. Cet effet stabilisant pourrait être le résultat d'un meilleur empaquetage du noyau hydrophobe au niveau des cystéines. Le Trp remplirait donc une partie du trou créé lors du remplacement de la Leu et de la Met par les cystéines, et ce, même s'il est exposé partiellement au solvant. Mais cette augmentation de structure n'est pas accompagnée par une augmentation de stabilité conformationnelle. La protéine TrpLH résiste moins bien que MB-1LH à la dénaturation thermique et protéolytique. Le tryptophane aiderait donc à spécifier la connectivité main gauche pour TrpLH mais ne contribuerait pas à augmenter sa stabilité conformationnelle. Ceci peut être expliqué par le fait que pour augmenter la stabilité thermique de MB-1LH il faut augmenter la surface hydrophobe cachée à l'intérieur de la protéine lorsque celle-ci est pliée. Or on sait que même si le tryptophane a une plus

grande surface hydrophobe que la tyrosine, ce tryptophane n'est pas bien isolé de l'eau limitant ainsi son effet de stabilisant thermique. De plus, l'insertion du Trp a créé une fluidification du noyau hydrophobe qui se détecte par la formation de niche(s) hydrophobe(s) à la surface de la protéine. Ces niches sont composées de résidus hydrophobes exposés au solvant, ce qui diminue d'autant plus la quantité de surface hydrophobe cachée à l'intérieur de la protéine. Ce modèle expliquerait pourquoi on observe une augmentation du taux de structures secondaires lors de l'insertion du Trp chez MB-1LH alors que l'on observe une perte de stabilité protéolytique et thermique.

Ce modèle peut aussi expliquer le comportement de MB-1TrpRH par rapport à MB-1RH. On constate que la protéine MB-1RH éprouve une perte de structures secondaires hélicales lors du changement de sa tyrosine 62 par un tryptophane alors que son pont S-S est intact. Cette perte de structure se reflète par une très grande fluidification du noyau hydrophobe et le tryptophane est au moins partiellement exposé au solvant. La perte de structure peut être due au fait que la position mutée n'est pas favorable à l'insertion d'un gros résidu lorsque la protéine adopte une conformation de connectivité main droite. De plus, le Trp n'amène aucune stabilisation de la protéine lorsque le pont S-S est ouvert. Il n'amène aucune augmentation du T_m mais amène une faible amélioration de la stabilité protéolytique. Le Trp ne remplit pas du tout son rôle visant à emplir le trou créé par l'insertion des cystéines. La mutation n'augmente pas la quantité de surface hydrophobe cachée dans le cœur de la protéine comme on s'y attendrait mais ne la diminue pas non plus à en juger par les expériences de dénaturation thermique. Intuitivement, on aurait pu s'attendre à une telle baisse compte tenu de la chute des structures secondaires hélicales. La protéine a donc trouvé un moyen de cacher les résidus hydrophobes qui se retrouvaient dans les segments d'hélices dépliés par l'insertion du Trp. Ce facteur doit sûrement contribuer à l'augmentation marquée de la présence de poches hydrophobes à la surface de la protéine. Le Trp n'est donc pas inséré dans le noyau comme il devrait l'être et les explications pour ce phénomène sont les mêmes que celles énumérées à la section 3.3.2. La faible hausse de stabilité protéolytique de TrpRH par rapport à MB-1RH n'est pas due à une augmentation de

stabilité conformationnelle. Selon moi, il s'agirait plutôt d'un effet de changement de conformation de la protéine qui rendrait inaccessible un ou plusieurs sites d'attaque des protéases. Ce changement de conformation pourrait provenir de la perte des structures hélicales lors de l'insertion du Trp dans MB-1RH. La protéine se réarrangerait pour cacher ses groupements hydrophobes ce qui pourrait masquer des sites potentiels de protéolyse.

3.3.4 MB-1TrpLH vs. MB-1TrpRH

Les résultats concernant ces deux mutants ont déjà été traités en profondeur dans la discussion de l'article scientifique présentée au Chapitre 2 ainsi que dans les autres sections de ce chapitre et je ne ferai donc que les résumer dans cette section. L'emphase sera mise sur la différence des résultats provenant de la différence de connectivité des deux protéines.

Nous voyons premièrement que la protéine TrpLH présente plus de structures hélicales que TrpRH lorsque le pont disulfure est formé. Cette différence ne se retrouve pas entre les protéines MB-1LH et MB-1RH, elle est donc directement attribuable au remplacement de la tyrosine 62 par un tryptophane chez ces protéines. On constate que la position 62 est sensible à la connectivité de la protéine lorsque celle-ci est occupée par un Trp. La sensibilité de cette position à la différence de connectivité pourrait s'expliquer par le fait que les différents résidus composant l'environnement du tryptophane 62 changent de position les uns par rapport aux autres lorsque l'on change la connectivité de la protéine. C'est justement à partir de ce principe qu'il est possible de spécifier la connectivité main droite ou gauche par l'insertion de cystéines en des positions spécifiques. Il n'est donc pas surprenant que ce changement d'environnement influence la façon dont le Trp peut s'insérer dans le cœur hydrophobe de la protéine. L'insertion d'un Trp au sein des protéines MB-1LH et RH confère donc des propriétés différentes à chacune des deux protéines.

Comme on le sait, deux différentes forces agissent au niveau de nos protéines à pont disulfure, soit l'effet stabilisant du pont et l'effet déstabilisant des cystéines. Ces deux effets sont mesurés en observant les caractéristiques des protéines lorsque le pont est fermé (mesure de l'effet du pont sur la protéine) et lorsqu'il est ouvert (mesure de l'effet des cystéines). Il a aussi été mentionné que l'insertion du Trp chez MB-1LH a apporté un gain de structure hélicales lorsque la protéine est réduite. Ceci signifie que l'insertion du Trp a diminué le caractère déstabilisant des cystéines chez TrpLH. Les cystéines ont un effet moins négatif sur la structure de cette protéine que le pont soit formé ou non, ce qui n'est pas le cas pour la protéine MB-1TrpRH. Pour ce qui est des autres paramètres, ils ont déjà été abordés dans les autres sections de la discussion.

CHAPITRE IV

CONCLUSION

Plusieurs résultats concernant les protéines MB-1TrpLH et MB-1TrpRH ont été présentés et discutés dans les sections précédentes. Ces résultats ont été placés dans le contexte des résultats obtenus pour d'autres protéines de la famille MB-1 et des interprétations et hypothèses ont été amenées. Quelles conclusions pouvons nous tirer de tous ces résultats ?

Premièrement, nous pouvons dire que l'effet de l'insertion d'un tryptophane en remplacement de la tyrosine 62 au sein des protéines MB-1LH et MB-1RH n'a pas donné les résultats attendus. Nous n'avons noté aucune augmentation de la température de dénaturation chez les mutants produits et l'effet inverse a même été observé pour MB-1TrpLH. Une faible augmentation de la stabilité protéolytique s'est manifesté chez TrpRH mais cette augmentation n'a pas réussi à rendre cette protéine plus stable que la protéine MB-1Trp. Ces résultats sont entre autres attribuables à une mauvaise insertion du Trp au sein du cœur hydrophobe chez les protéines à pont disulfure. Différentes solutions pourraient être envisagées afin de combiner la présence d'un Trp et de pont disulfures à l'intérieur de MB-1. Une des solutions serait de déplacer les cystéines dans les connecteurs qui relient les hélices les une aux autres. Cette modification ne devrait pas amener l'effet déstabilisant qui a été observé lorsque l'on a inséré les cystéines dans le cœur hydrophobe de la protéine. Toutefois, si les cystéines n'occupent plus des positions précises au sein des hélices, il sera impossible de spécifier la topologie main droite ou gauche de celles-ci. L'insertion d'un ou plusieurs ponts disulfures dans les régions non structurées de la protéine n'amènerait pas beaucoup d'informations fondamentales au niveau de l'étude de la structure protéique mais pourrait tout de même conférer assez de

stabilité à la protéine pour nous permettre de l'utiliser pour l'alimentation des vaches laitières.

L'étude des mutants TrpLH et TrpRH nous permet aussi d'affirmer que le taux de structures secondaires d'une protéine n'est pas corrélé à la stabilité thermique et protéolytique de la protéine. Ainsi, malgré le fait que TrpRH présente moins de structures hélicales que TrpLH, il résiste mieux à la dénaturation et à la protéolyse. Le remaniement de la protéine MB-1 dans le but de lui spécifier une seule et unique topologie pourrait amener la protéine à adopter une structure se rapprochant plus d'une protéine native mais ne garanti en aucun cas une augmentation de stabilité thermique ou protéolytique. Ceci est dû au fait que les facteurs amenant une augmentation de stabilité chez une protéine ne sont pas nécessairement les mêmes que ceux qui spécifient la topologie de cette même protéine.

On peut faire une corrélation très nette entre la température de dénaturation des protéines de la famille MB-1 et leur susceptibilité à subir la protéolyse. Ainsi, les deux mutants présentant les T_m les plus faibles sont aussi ceux qui résistent le moins à la protéolyse, alors que MB-1Trp présente le T_m le plus élevé et est aussi la protéine qui résiste le mieux aux protéases. Pour ce qui est des mutants présentant des températures de dénaturation intermédiaires, il faut analyser plus en détail leur courbe de dénaturation thermique. Ainsi, deux protéines ayant le même T_m peuvent présenter des courbes de dénaturation thermique différentes. Elles présenteront donc des fractions de protéines dépliées différentes pour une même température (autre que le T_m évidemment). C'est pour cette raison que les différents mutants de MB-1 ayant des T_m intermédiaires (près de 50°C) ont des stabilités protéolytiques légèrement différentes les unes des autres. La cause d'une telle corrélation entre le T_m et la protéolyse est évidente. Une protéine pliée et compacte offre beaucoup moins de sites d'attaque aux protéases. Par contre, une protéine partiellement dépliée à la même température présente des parties totalement exposées au solvant et donc facilement accessibles aux protéases. À partir de cette

corrélation nous pouvons donc concentrer nos efforts à améliorer seulement l'un des deux paramètres chez MB-1 tout en sachant qu'un l'autre paramètre réagira dans le même sens.

Finalement, à la lumière des résultats obtenus pour mes deux protéines ainsi que ceux obtenus pour les autres protéines de la famille MB-1, nous pouvons affirmer qu'il est très difficile de prédire l'effet d'une mutation chez MB-1 ou ses homologues. Prenons par exemple le remplacement de la Tyr 62 par un Trp. Cette mutation a été effectuée chez MB-1, MB-1LH et MB-1RH et les effets de cette mutation sont différents chez les trois protéines pourtant très semblables au niveau de leur séquence en acides aminés. Il en va de même pour ce qui est de l'insertion de ponts disulfures chez MB-1 et MB-1Trp. Les changements engendrés par ces mutations sont différents pour les quatre protéines générées. L'effet déstabilisant des cystéines est cependant présent chez toutes les protéines construites. La fermeture du pont contrebalance aussi l'effet déstabilisant dû à l'insertion des cystéines et ce, pour toutes les protéines. Dans ce cas, c'est l'intensité de la déstabilisation engendrée par les cystéines et de la stabilisation lors de la fermeture du pont qui est différent pour les quatre mutants. En ce qui concerne l'insertion de ponts disulfures au sein de protéines naturelles, les résultats obtenus sont plutôt variés (23, 24, 25). Certains amènent une stabilisation alors que d'autres déstabiliseront la protéine. Les résultats quelque peu disparates que j'ai obtenus vont donc dans le même sens que la littérature.

CHAPITRE V

OUVERTURE DE RECHERCHES FUTURES

Plusieurs ouvertures de recherche se présentent à nous pour atteindre nos buts. On peut développer la branche fondamentale du projet qui est de déterminer la relation entre la séquence de la protéine à sa structure. On peut aussi axer le projet dans un but plus appliqué qui est la production de protéines stables et enrichies en A.A.E. destinées à la production animale. Plusieurs méthodes peuvent être utilisées pour parvenir à nos fins.

5.1 Spécification de la connectivité

Du point de vue fondamental, on pourrait tenter de spécifier la connectivité de MB-1 en utilisant des spécificateurs autres que les ponts disulfures. Ce pourrait être par la formation de ponts salins disposés un peu partout au sein de la protéine. Ces ponts salins sont retrouvés chez les protéines naturelles et sont épars, contrairement aux fortes concentrations de résidus chargés retrouvés chez les peptides synthétiques généralement utilisés en ingénierie des protéines. La spécification de la topologie pourrait aussi être effectuée par un empaquetage spécifique du cœur hydrophobe ne permettant qu'une seule des deux connectivités. Ces options présentent toutefois un désavantage certain. Il faudrait effectuer une panoplie de mutations au sein de la protéine et il est difficile de prédire l'effet d'une mutation unique sur celle-ci donc encore plus difficile de prédire l'effet de plusieurs mutations combinées. D'autres méthodes peuvent être utilisées pour atteindre nos buts. L'évolution accélérée est l'une d'elle.

5.2 Évolution accélérée et sélection à haut volume

L'évolution accélérée est une méthode par laquelle on mute un gène ou une partie de gène de façon tout à fait aléatoire. Il existe plusieurs techniques pour muter un gène de façon aléatoire. On peut nommer la polymérisation en chaîne à haut taux d'erreurs (*error-prone PCR*), l'utilisation de cassette mutagénique dont la séquence a été choisie au hasard, la mutation par oligonucléotides de séquences choisies au hasard, l'utilisation de souches bactériennes induisant des erreurs au niveau de l'ADN, etc. Toutes ces techniques permettent de créer une banque de mutants composés d'une multitude de séquences différentes du gène ciblé. Par exemple, si l'on mute en même temps au hasard cinq acides aminés sur une séquence de 200 acides aminés totaux, on pourra générer 6 278 520 528 393 960 séquences différentes (26). La plupart de ces séquences ne coderont pas pour une protéine entière et capable de se replier convenablement. Il faut donc absolument être en mesure de sélectionner de façon très rapide les quelques séquences codant pour une protéine stable et de pleine longueur. Des techniques de sélection à haut volume (*high throughput screening*) doivent donc être mises au point. La technique que notre laboratoire se prépare à utiliser implique l'expression de protéines à la surface d'un bactériophage. La protéine à être exprimée doit présenter un site ayant une certaine affinité pour un substrat quelconque à son extrémité. On fixe ce substrat à un support et on applique les phages exprimant une protéine à sa surface. Si la protéine reconnaît le substrat, elle (ainsi que le phage auquel elle est rattachée) restera fixée au support. Il est donc possible de mettre au point une banque de phages qui se compose de particules se distinguant par la séquence du gène de la protéine mutée au hasard et destinée à être exprimée à la surface du phage. De ces millions de clones possibles, il sera aisé de faire ressortir ceux présentant une protéine complète puisque seulement les protéines complètes présenteront le site de fixation pour le substrat particulier. En utilisant cette méthode, on peut aussi modifier les conditions de sélection de façon à enrichir seulement les protéines avec un caractère particulier. Dans notre cas, on envisage d'exposer les phages présentateurs de protéines à un mélange de protéases avant la sélection par affinité. Cette technique devrait permettre d'éliminer les phages

présentant des protéines incomplètes ainsi que ceux présentant une protéine sensible à la protéolyse. De cette façon, nous pourrions repêcher les quelques séquences codant pour des protéines stables parmi les millions de séquences possibles et ce dans un laps de temps relativement court. Le seul problème envisagé lors de ces manipulations est la diminution du taux d'A.A.E. chez les protéines sélectionnées comme étant stable. Il sera donc important d'ajuster l'astringence de la sélection afin de permettre un bon équilibre entre un taux d'A.A.E. maximum et la stabilité de la protéine. L'autre facteur qui sera négligé lors de cette étude est la mise en lumière de la relation entre la séquence protéique et la structure de cette même protéine. Comme les mutations seront multiples, il sera presque impossible de déterminer l'effet de chacune des mutations individuellement et de son apport à la stabilité de la protéine. Toutefois, cette méthode sera très utile pour produire des protéines stables à haute valeur nutritive.

5.3 Cristallographie et modélisation

Une dernière solution pour tenter d'améliorer la stabilité de la protéine MB-1 serait d'obtenir la structure cristallographique du mutant le plus stable, soit MB-1 Trp. Si l'on avait cette structure, on serait capable de faire de la modélisation par ordinateur. On pourrait alors choisir des mutations de façon rationnelle et ainsi remplir correctement le cœur hydrophobe, favoriser la formation de ponts salins, de ponts disulfures, etc. Ceci nous permettrait d'améliorer la stabilité de la protéine et aussi de spécifier sa connectivité. Des essais de cristallisation ont été tentés avec MB-1 mais la structure de la protéine n'a pu être déterminée. Tenter les mêmes essais avec une protéine plus stable pourrait éventuellement s'avérer fructueux mais ces tests restent à être effectués.

BIBLIOGRAPHIE

-
- 1 Voet , D., Voet, J. G., “Biochemistry”, 2e éd., John Wiley & Sons, Inc., U.S.A., pp. 106 (1995).
 - 2 Voet , D., Voet, J. G., “Biochemistry”, 2e éd., John Wiley & Sons, Inc., U.S.A., pp. 57 (1995).
 - 3 Voet , D., Voet, J. G., “Biochemistry”, 2e éd., John Wiley & Sons, Inc., U.S.A., pp. 141-190 (1995).
 - 4 Garrett, R. H., Grisham, C. M., “Biochimie”, 2e éd., De Boeck Université, France, pp. 158-208 (2000).
 - 5 Ratnaparkhi, G. S., Varadarajan, R., “Thermodynamic and structural studies of cavity formation in proteins suggest that loss of packing interaction rather than hydrophobic effect dominates the observed energetics”, *Biochemistry* 39 : 12365-12374 (2000).
 - 6 Pace, C. N., Shirley, B. A., Thomson, J. A., “Protein structure: A practical approach”, Creighton T. E ed., IRL Press, Oxford, England, pp. 311-330 (1989).
 - 7 Creighton, T. E., “Proteins : Structures and Molecular Properties”, W. H. Freeman, New York, U.S.A., pp. 133-157 (1983).
 - 8 Tinoco, I. Jr., Sauer, K., Wang, J. C., “Physical Chemistry : Principles and applications in Biological Sciences”, 3e éd., Prentice-Hall, Inc. New Jersey, U.S.A., pp. 67-122, (1985).
 - 9 Voet , D., Voet, J. G., “Biochemistry”, 2e éd., John Wiley & Sons, Inc., U.S.A., pp. 769 (1995).
 - 10 Munneke, R. L., Schingoethe, D. J., Casper, D. P., “Lactational avaluation of ruminally protected methionine in diets containing extruded soybeans and urea”, *J. Dairy Sci.* 74 : 227 (1991).
 - 11 King, K. J., Huber, J. T., Sadik, M., Bergen, W. G., Grant, A. L., King, V. L., “Influence of dietary proteins sources on the amino acids profiles available for digestion and metabolism in lactation cows”, *J. Dairy Sci.* 73 : 3208-3216 (1990).
 - 12 Hungate, R. E., “The rumen and its microbes” Academic press Inc., New York, U.S.A. (1966).

- 13 Aventis animal nutrition, "The Smartamine™ line of protected amino acids for ruminants", Uniform Resource Location : <http://www.an.aventis.com/public/a/cor96str3/>, (2001).
- 14 DeLumen, B. O., Krenz, D. C., Revilleza, J., "Molecular strategies to improve the protein quality of legumes", *Food Technology* 51 : 67-70 (1997).
- 15 Hoffman, L. M., Donaldson, D. D., Herman, E. M., "A Modified storage protein is synthesized, processed, and degraded in the seed of transgenic plants", *Plant Molecular Biology* 11 : 717-729 (1988).
- 16 Beauregard, M., Dupont, C., Teather, R. M., Hefford, M. A., "Design, expression, and initial characterization of MB1, a *De Novo* protein enriched in essential amino acids", *Biotechnology* 13 : 974-981 (1995).
- 17 Hefford, M. A., Dupont, C., MacCallum, J., Parker, M. H., Beauregard, M., "Characterization of MB-1: A dimeric helical protein with a compact core", *Eur. J. Biochem.* 262 : 467-474 (1999).
- 18 MacCallum, J., Hefford, M. A., Omar, S., Beauregard, M., "Prediction of folding stability and degradability of the de novo designed protein MB-1 in cow rumen", *Appl. Biochem. Biotechnol* 66 : 83-93 (1997).
- 19 Wedemeyer, W. J., Welker, E., Narayan, M., Scheraga, H. A., "Disulfide bonds and protein folding", *Biochemistry* 39 : 4207-4216 (2000).
- 20 Gagnon, M. C., Williams, M., Doucet, A., Beauregard, M., "Replacement of Tyr62 by Trp in the designer protein Milk Bundle-1 results in significant improvement of conformational stability", *FEBS Letters* 484 : 144-148 (2000).
- 21 Betz, S. F., Liebman, P. A., DeGrado, W. F. "De novo design of native proteins: Characterization of proteins intended to fold into antiparallel, Rop-like, four helix bundle", *Biochemistry* 36 : 2450-2458 (1997).
- 22 Parsell, D. A., Sauer, R. T., "The structural stability of a protein is an important determinant of its proteolytic susceptibility in *Escherichia coli*", *J. Biol. Chem.* 264 : 7590-7595 (1989).
- 23 Betz, S. F., "Disulfide bonds and the stability of globular proteins", *Protein Sci.* 2 : 1551-1558 (1993).
- 24 Hinck, A. P., Truckses, D. M., Marley, J. L., "Engineering disulfide bonds in staphylococcal nuclease: effect of the stability and conformation of the folded protein", *Biochemistry* 35 : 10328-10338 (1996).

-
- 25 Robinson, C. F., Sauer, R. T., "Striking stabilization of Arc repressor by engineering disulfide bonds", *Biochemistry* 39 : 12494-12505 (2000).
- 26 Kuchner, O., Arnold, F. H., "Directed evolution of enzyme catalysts", *Trends Biotechnol.* 15 : 523-530 (1997).

ANNEXE A

Trois-Rivières, le 24 juillet 2001

Objet : demande de rédaction de mémoire de maîtrise sous forme d'article scientifique.

Demande faite à M. Pierre Tancrede, directeur du Comité d'études avancées en Biophysique et Biologie cellulaire.

Par la présente, je voudrais vous demander l'autorisation de déposer mon mémoire de maîtrise sous forme d'article scientifique.

Le titre de mon mémoire sera Combinaison de deux types de mutations stabilisantes au sein de la protéine de novo MB-1. L'article que je souhaite utiliser comme base pour la rédaction de mon mémoire s'intitule *Engineering Nutritious Proteins : Improvement of Stability in the Designer Protein MB-1 via introduction of disulphide bridges*. Cet article a été soumis au journal scientifique Journal of Agricultural and Food Chemistry et la confirmation de la réception par l'éditeur est jointe à cette lettre. Je suis l'auteur principal de cet article rédigé en anglais. Les autres parties du mémoire seront, quant à elles, rédigées en français.

Je joins aussi à cette lettre une lettre expliquant ma contribution et celle des autres auteurs à l'article scientifique mentionné précédemment.

Veuillez agréer mes salutations les plus sincères

Alain Doucet

Étudiant à la maîtrise en biophysique.

Alain_doucet@uqtr.quebec.ca

1022 pavillon Pierre-Boucher, U.Q.T.R.

(819) 376-5011 #3327

Contribution des différents auteurs

Contribution des différents auteurs à l'article intitulé Engineering Nutritious Proteins : Improvement of stability in the Designer Protein MB-1 via Introduction of Disulfide Bridges.

L'article présente quatre mutants de la protéine MB-1. J'ai fait ou dirigé toutes les expérimentations et manipulations visant deux des quatre mutants. J'ai aussi fait l'analyse et l'interprétation des résultats concernant ces deux protéines. Les deux autres protéines citées dans la manuscrit ont été produites par M. Martin Williams (2^{ième} auteur de l'article). J'ai aussi participé aux mesures de dichroïsme circulaire visant les deux protéines de M. Williams.

En ce qui a trait à la rédaction de l'article, j'ai écrit les sections Introduction, Experimental Procedures et Results et produit toutes les figures s'y rattachant. J'ai aussi eu une participation active pour la rédaction des autres sections de l'article.

Pour ce qui est des autres auteurs, M. Martin Williams a produit les deux autres protéines citées dans l'article et a fourni les résultats et interprétation pour ces deux mutants. Mme. Mylène Gagnon a fourni un support lors de la production des protéines et participé à des discussions traitant de l'interprétation des données. Finalement, M. Marc Beauregard a supervisé et financé tout le projet.

Globalement, j'estime avoir contribué pour plus de 60% à l'article scientifique dont il est question.

Cordiales salutations

Alain Doucet

Étudiant à la maîtrise en biophysique



PUBLISHED BY
THE AMERICAN CHEMICAL SOCIETY

Kraft Foods
801 Waukegan Road
Glenview, IL 60025
Phone: (847) 646-7321
Fax: (847) 646-3301
E-mail: jfinley@kraft.com

John W. Finley, *Associate Editor*

August 30, 2001

Dr. Marc Beauregard
Dept. de Chimie-Biologie
Universite du quebec a Trois-Rivieres
3357 Des Forges.
Trois-Rivieres QC G9A 5H7
CANADA

RE: Manuscript #JF010839D entitled "Engineering Nutritious Proteins: Improvement of Stability in the Designer Protein MB-1 Via Introduction of Disulfide Bridges"

Dear Dr. Beauregard:

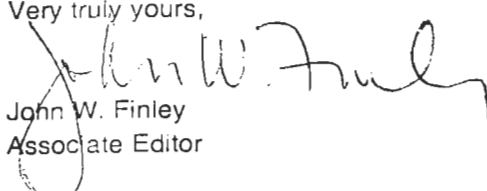
Reviews of your manuscript have been received and are enclosed. I am pleased to accept your manuscript for publication subject to your revision in response to comments made by the reviewers. You should give careful consideration to these comments and revise the manuscript as may be appropriate.

Please follow the Guidelines for Authors very carefully, as it will lengthen publication time if additional corrections are required. PLEASE PAY CAREFUL ATTENTION TO THE NEW REFERENCE STYLE UNDER "LITERATURE CITED". As of January 1, 2001 we ask that you list your references numbered in the order that they appear in the text.

Please return two hard copies of the revised manuscript along with an itemized list of the changes that have been made for use by my office and an identical version of the revised manuscript on disk. Unfortunately, we are not able to use graphics from the disk, so you must provide a high quality hard copy of each figure for photographic reproduction. Also, a title running-header is required for publication.

It is our goal to minimize the time for publication, and we request your help by returning all of these items within 4 weeks. Please let us know if you anticipate any significant delay or problems with providing a suitable version on disk so that we can plan our publication schedule accordingly.

Very truly yours,


John W. Finley
Associate Editor

Enclosures

Scope and Editorial Policy

Journal Coverage. The *Journal of Agricultural and Food Chemistry* publishes fundamental and applied research papers in a broad scientific area dealing with the chemistry and biochemistry of all aspects of agriculture and food. The Journal also encourages papers combining chemistry and/or biochemistry with biological/sensory/toxicological evaluation. Contributed papers must report original research, that is, have a definable impact on the advancement of science and technology. Originality will be documented by novel experimental results, theoretical treatments, and interpretations of data. The Journal is organized into the following sections: Food Chemistry/Biochemistry; Composition of Foods/Feeds/Biobased Products; Biotechnology; Nutritional Activity; Bioactive Constituents; Safety/Toxicology; Environmental Chemistry; Flavors and Aromas; Chemical Changes during Processing/Storage; Chemistry of Crop and Animal Protection; and Analytical Methods. The Editors will periodically evaluate these classifications to ensure that they remain representative of areas encompassed by the Journal.

Unsolicited review articles are considered which summarize information in a field in which the literature is scattered, or treat published data or other information so as to provide a new approach, or stimulate further research. Authors considering the preparation of a review should submit a synopsis to the Editor to establish whether the manuscript will meet these guidelines. The Editors may occasionally invite manuscripts that review particularly timely or noteworthy areas of agricultural and food chemistry. The Journal will also consider publication of a series of manuscripts reporting original research that are presented in a symposium or otherwise clustered around a single topic. Prospective organizers should contact the Editor well in advance to determine whether the subject matter conforms to the Journal's goals, criteria, and available space. There is no separate Rapid Communications or Notes section. All manuscripts, even relatively brief ones, will be considered for publication based upon the Journal's criteria.

Comments related to published papers will be considered from readers if the correspondence is received within six months of the date of publication of the original paper; the authors of the original paper will be given the opportunity to reply to such comments within two months, if they so desire. Both comments and replies should not exceed 1000 words each, including citations, and will be published consecutively in the same issue of the Journal.

Submission of Manuscripts. Four copies of manuscripts should be submitted to

James N. Seiber, Editor

Journal of Agricultural and Food Chemistry

Department of Environmental Toxicology

University of California, Davis

One Shields Avenue

Davis, California 95616

Manuscripts must be written in English, printed double-spaced in a standard font on 22 × 28 cm (or A4) paper, with all pages and lines of the manuscript,

including tables and figures, numbered. For specific details regarding format, see Manuscript Format in the accompanying Preparation of Manuscripts and refer to a current issue for general style. Preparation of Manuscripts guidelines are published in the first issue of each volume and are also available at <http://pubs.acs.org/journals/jafcau/index.html>. Authors should also familiarize themselves with the Ethical Guidelines following these instructions or at <http://pubs.acs.org/instruction/ethic.html>. The *ACS Style Guide* (2nd ed., 1997; ISBN 0-8412-3462-0), available from Oxford University Press, Order Department, 201 Evans Road, Cary, NC 27513, provides a detailed treatment of the fundamentals of manuscript preparation.

The manuscript should be accompanied by a covering letter indicating the particular significance and originality of the research, the relevance of the subject matter to research in agricultural and food chemistry, and the names/addresses of four suggested reviewers. Authors should state reasons if expedited review and processing of the manuscript are requested.

Editorial Procedures and Peer Review. To ensure the highest possible quality, all manuscripts submitted are subject to peer review. Scientists with expertise in the subject matter being treated will evaluate the manuscript for validity of the experimental design and results, originality, significance, and appropriateness to the Journal. Authors who are not familiar with idiomatic English are urged to seek assistance; deficiencies in grammar may be a serious hindrance during the review process. The Editors may exercise their prerogative to reject a manuscript without peer review if the paper is judged to be outside the scope of the Journal, poorly written or formatted, fragmentary and marginally incremental, or lacking in significance.

All manuscripts submitted are reviewed and handled by the Editor-in-Chief or assigned to one of the four Associate Editors. The Associate Editor and local Editorial Assistant are then responsible for the assigned manuscripts, including acknowledging receipt, evaluating the content of the paper, selecting reviewers, monitoring the progress of the review process, evaluating the comments of reviewers and forwarding them to the authors for their response, communicating ultimate acceptance or rejection to the corresponding author, and carrying out a final check of accepted manuscript for appropriate format and style. Typically, two or three reviewers are selected per paper on the basis of the subject matter, available expertise, and the Editor's knowledge of the field. The obligations of Editors and Reviewers are outlined in the Ethical Guidelines.

Potential reviewers for each paper are identified by various means, including a computerized search of the subject area. Authors must submit the names and addresses (including e-mail addresses and fax numbers) of four potential reviewers; however, the Editors are under no obligation to use specific individuals. Reviewers are normally asked to provide their assessments within two to three weeks. Anonymous copies of reviews and the Editor's decision regarding the acceptability of the manuscript are sent to the corresponding author. If the reviewers' evaluations of the manuscript

disagree, or if reviewer's and Editor's comments are not satisfactorily addressed by the authors, the Editor may reject the manuscript or select additional reviewers. These additional reviews are used by the Editor to reach the final decision about the disposition of the manuscript.

Notice: Documents accepted for publication in ACS Journals will be posted in the World Wide Web edition of the journal as soon as they are ready for publication, that is, when galley proofs are corrected and all author concerns are resolved. This can occur anywhere from 2 to 11 weeks in advance of the cover date of the printed issue. Authors should take this into account when planning their intellectual and patent activities related to a document. The actual date on which the document is posted on the Web is recorded in a separate line at the bottom of the first page of the document in the printed issue.

Conflict of Interest. Research involving an evaluation of commercial products should not reveal the brand names of such products unless information regarding their manufacture has been provided for publication by the company producing the product.

Codes such as letters (A, B, C, etc.) or numbers (1, 2, 3, etc., or I, II, III, etc.) may be used for purposes of identification. If brand names are used, the authors should disclose at the time of submission any financial arrangement they may have with a company whose product figures prominently in the submitted manuscript or with a company making a competing product. An editorial decision will then be made as to whether the manuscript being submitted should be sent out for review. If the paper is deemed suitable for review, information concerning any financial arrangement the authors may have with a given company will be held in confidence and will not influence the evaluation of the scientific merits of the research and whether the manuscript can be accepted for publication. As a general guiding principle, however, it is expected that the authors of such papers should not have any financial interest in a company (or its competitor) that makes a product discussed in the paper. These guidelines do not generally apply to the use of brand names or to the identification of the producers of products that are used for analytical purposes such as instruments, reagents, or kits.

Preparation of Manuscripts

(Revised January 2001)

Manuscript Format. Manuscripts must be prepared with a typewriter or word processor and be printed double-spaced, including abstracts, references, tables, and figure captions. All pages must be numbered consecutively starting with the title page and including tables and figures. The Journal requires authors to stay within a 20 typed page limit, not including references, tables, and figures. The various sections of the manuscript should be assembled in the following sequence:

- Title and Authorship (single page)
- Abstract and Keywords (single page)
- Introduction
- Materials and Methods
- Results
- Discussion
- Abbreviations Used
- Safety
- Acknowledgment
- Supporting Information Description
- Literature Cited
- Figure Captions
- Tables
- Figures

The title, authorship, and institutional affiliations should be included on a single page. Use specific and informative titles. Keywords in the title assist in effective literature retrieval. If a plant is referred to in the title or elsewhere in the text by its common or trivial name, it should be identified by its scientific name in parentheses immediately following its use. This term should also be provided as one of the keywords. If trade names are mentioned, give generic names in parentheses. The titles of all papers will appear in the Table of Contents of each issue under specific subject headings. Authors will be given the opportunity of selecting the most appropriate subheading from a list provided to them prior to publication.

Authorship. Be consistent in authorship designation. First name, middle initial, and last name are generally adequate for correct identification, but omit titles. Give the complete mailing address of the institution where the work was conducted and identify the affiliation of each author. If the current address of an author is different, include it in a footnote on the title page. The name of the author to whom inquiries about the paper should be addressed must be marked with an asterisk; provide the telephone and fax numbers and e-mail address of this correspondent.

Abstract. Authors' abstracts are now used directly for *Chemical Abstracts*. The abstract should be a clear, concise (100–150 words) one-paragraph summary, informative rather than descriptive, giving scope and purpose, experimental approach, significant results, and major conclusions. Write for literature searchers as well as journal readers.

Keywords. Provide significant keywords to aid the reader in literature retrieval. The keywords are published immediately before the text, following the abstract.

Introduction. The purpose or reason for the research being reported, and its significance, should be clearly and concisely stated. Discuss relationships of the study

to previously published work, but do not reiterate or attempt to provide a complete literature survey.

Materials and Methods. Apparatus, reagents, and biological materials used in the study should be incorporated into a general section. List devices of a specialized nature or instruments that may vary in performance, such that the model used may affect the quality of the data obtained (e.g., spectroscopic resolution). List and describe preparation of special reagents only. Reagents normally found in the laboratory and preparations described in standard handbooks or texts should not be listed. Specify the source, vendor [city and state (or country)], and availability of special equipment, reagents, kits, etc.

Biological materials should be identified by scientific name (genus, species, authority, and family) and cultivar, if appropriate, together with the site from which the samples were obtained. Specimens obtained from a natural habitat should be preserved by deposit of samples in an herbarium for plants or in a culture collection for microorganisms, with a corresponding collection or strain number listed.

Specific experimental methods should be sufficiently detailed for others to repeat the experiments unequivocally. Omit details of procedures that are common knowledge to those in the field. Brief highlights of published procedures may be included, but details must be left to literature cited. Describe pertinent and critical factors involved in reactions so the method can be reproduced, but avoid excessive description. For information on the reporting of certain types of data see the section below on Reporting Specific Data.

Results and Discussion. Results and discussion may be presented in separate sections or combined into a single section, whichever format conveys the results in the most lucid fashion. Be complete but concise in discussing findings, comparing results with previous work and proposing explanations for the results observed. Avoid comparisons or contrasts that are not pertinent and avoid speculation unsupported by the data obtained. A separate summary or conclusion section is not to be used; any concluding statements are to be incorporated under Results and Discussion.

Tables and figures should be carefully designed to maximize presentation and comprehension of the experimental data with superfluous information excluded. Useful information not directly relevant to the discussion may be included under Supporting Information (see below). Authors should note that structural formulas are valuable in expressing concisely the precise nature of the compounds under discussion and revealing the essence of the subject to readers unfamiliar with the topic, without their necessary recourse to reference materials. The use of chemical names without accompanying structures is little more than a narrative account and may cause readers to overlook the significance of the paper.

Abbreviations Used and Nomenclature. Standard abbreviations, without periods, should be used throughout the manuscript. Refer to *The ACS Style Guide* for the preferred forms of commonly used abbreviations. Specialized abbreviations may be used provided they are placed in parentheses after the word(s) for which they are to substitute at first point of use and are again defined in this Section. Avoid trivial names and "code

abbreviations (e.g., NAR for naringenin) unless such codes are in common usage (e.g., MTBE for methyl *tert*-butyl ether). If trade names are used, define at point of first use. If nomenclature is specialized, include a "Nomenclature" section at the end of the paper, giving definitions and dimensions for all terms. Use SI units insofar as possible.

Write all equations and formulas clearly and number equations consecutively. Place superscripts and subscripts accurately; avoid superscripts that may be confused with exponents. Identify typed letters and numbers that might be misinterpreted, such as "oh" for zero or "el" for one.

It is the authors' responsibility to provide correct nomenclature. All nomenclature must be consistent and unambiguous and should conform with current American usage. Insofar as possible, authors should use systematic names similar to those used by Chemical Abstracts Service, the International Union of Pure and Applied Chemistry, and the International Union of Biochemistry and Molecular Biology.

Chemical Abstracts (CA) nomenclature rules are described in Appendix IV of the *Chemical Abstracts Index Guide*. For CA nomenclature advice, consult the Manager of Nomenclature Services, Chemical Abstracts Service, P.O. Box 3012, Columbus, OH 43210-0012. A name generation service is available for a fee through CAS Client Services, 2540 Olentangy River Road, P.O. Box 3343, Columbus, OH 43210-0334; Telephone: (614) 447-3870; Telefax: (614) 447-3747; e-mail: answers@cas.org.

Safety. Authors are required to call special attention, in both their manuscripts and their covering letter, to safety considerations such as explosive tendencies, special precautionary handling procedures, and toxicity.

Acknowledgment. Include essential credits but hold to an absolute minimum; omit academic and social titles. Meeting presentation data and acknowledgment of financial support of the work should be included in a note following the Literature Cited.

Literature Cited. References should be listed on a separate sheet and numbered in the order in which they are cited in the text. Authors are responsible for the accuracy of their references. Give complete information, using last name and initials of the author, patentee, or equivalent; do not use "Anonymous". Follow *Chemical Abstracts Service Source Index* for abbreviations of journal titles. Because subscribers to the Web edition of the Journal are now able to click on the "CAS" tag following each reference to retrieve the corresponding abstract from *Chemical Abstracts*, **reference accuracy is critical**. References should be cited in the text by an on-line italic number in parentheses, for example, (1). Consult *The ACS Style Guide* and current issues of the Journal for style and format. Typical references follow the styles given below:

For journals:

1. Brown, J.; Jones, M.; Green, D. Article title. *J. Agric. Food Chem.* 1980, 28, 1-4.

For books:

2. Smith, L.; Caldwell, A. Chapter title. In *Book Title*; Keys, F., Park, G., Eds.; Publisher: City, State (or Country), Year; Pages.

Papers should not depend for their usefulness on unpublished material, and excessive reference to material "in press" is discouraged. Reference to the authors'

own unpublished work is permitted if the subject is of peripheral relevance to the manuscript in question, but any unpublished results of central importance must be described in sufficient detail within the manuscript. If pertinent references are "in press" or unpublished for any reason, **furnish copies to enable reviewers to evaluate the manuscript.**

Figure Captions. Figures should be numbered consecutively with Arabic numerals. The accompanying caption should be concise but sufficiently descriptive for the figure to be understandable without recourse to the text. Blocks of chemical structures may be included in figures or in the text at a position as close as possible to the point of first reference.

Tables. These should be numbered consecutively with Arabic numerals and should be grouped at the end of the paper. Footnotes in tables should be given letter designations and be cited in the table by italic superscript letters. The sequence of letters should proceed by row rather than by column. Each table should be provided with a descriptive heading, which, together with the individual column headings, should make the table, as nearly as possible, self-explanatory. In setting up tabulations, authors are requested to keep in mind the type area of the journal page (17.8 × 25.4 cm), and the column width (8.5 cm), and to make tables conform to the limitations of these dimensions. Arrangements that leave many columns partially filled or that contain much blank space should be avoided. Conversely, arrangements that include >20 columns should be broken into two tables if possible.

Figures. Submit four sets of original drawings, or high-quality glossy photographic prints, of original graphs, diagrams, and structural formulas; photocopies do not reproduce well. If gel patterns are submitted, provide four sets of photographs of the original gel patterns. Photographs that are reproduced by a copier frequently do not provide sufficient resolution of individual bands. Do not exceed 8.5 × 11 in. (22 × 28 cm) dimensions. The quality of the illustrations printed in a paper largely depends on the quality of the illustrations provided. ACS journal pages are now produced completely electronically. The chemical structures, graphs, photographs, or other illustrations sent with a paper will be scanned into the journal page using a digital scanner. The scanner is very sensitive; it will faithfully copy all flaws such as smudges, uneven lines, and incomplete erasures.

In illustrations, contrast is important. Use dark black ink on high-quality, smooth, opaque white paper. Ordinary white bond paper works well. Avoid tracing paper or textured "artist" papers. A brochure providing more detailed information regarding guidelines for illustrations is also available from the editorial office upon request. Each figure or photograph should be properly identified by printing the figure number and author(s) on the reverse side, but only in the margin and not in the area of the figure itself because these marks may show through when the photograph is scanned.

Illustrations must fit a one- or two-column format on the journal page. **For efficient use of journal space, single-column illustrations are preferred.**

	single (preferred)	double
width		
minimum		10.5 cm (4.13 in.)
maximum	8.25 cm (3.25 in.)	17.78 cm (7 in.)
maximum depth	24 cm (9.5 in.)	24 cm (9.5 in.)

For best results, submit illustrations in the actual size at which they should appear in the journal. Original illustrations that do not need to be reduced to fit a single or double column will yield the best quality. Lettering should be no smaller than 4.5 points. (Helvetica type works well for lettering.) Lines should be no thinner than 0.5 point. Lettering and lines should be of uniform density.

If you must submit artwork that must be reduced, use larger lettering and thicker lines so that, when reduced, the artwork meets the above-mentioned parameters.

Avoid using complex textures and shading to achieve a three-dimensional effect. To show a pattern, choose a simple crosshatch design.

Artwork produced by a high-quality graphics plotter will yield a better printed result than artwork produced on a dot-matrix device, laser printer, or inkjet printer. If your material must be prepared on a dot-matrix, laser, or inkjet printer, use high-quality laser paper and choose the highest resolution available. For example, select 600 dpi rather than 300 dpi.

Photographs should be high-contrast prints with a smooth or glossy finish. Send photographs that are single- or double-column width so that they will not have to be reduced. Do not submit negatives, slides, or overhead transparencies. Avoid photographs produced on a laser printer and prints cut from a printed publication; these do not give good results.

Color photographs and artwork may be printed in the journal if the Editor approves of their use. **Color illustrations should be submitted only if they are essential for clarity of communication.** Reproduction of color illustrations will be provided at no cost to the author; however, a surcharge of \$100 per 100 reprints will be added to the standard cost of reprints. Do not submit color prints to be printed in black and white.

Structures should be produced with the use of a drawing program such as ChemDraw. Authors using the current versions of ChemDraw will find the necessary parameters incorporated into these programs ("ACS Document 1996").

Authors using older versions of ChemDraw should use the following parameters:

1. As drawing settings select:

chain angle	120°
bond spacing	18% of width
fixed length	14.4 points (0.508 cm, 0.2 in.)
bold width	2.0 points (0.071 cm, 0.0278 in.)
line width	0.6 point (0.021 cm, 0.0084 in.)
margin width	1.6 points (0.056 cm, 0.0222 in.)
hash spacing	2.5 points (0.088 cm, 0.0347 in.)

2. As text setting select:

font	Helvetica
size	10 pt

3. Under the preferences choose:

units	points
tolerances	3 pixels

4. Set the page setup to 100%.

5. Using the ChemDraw ruler or appropriate margin settings, create structure blocks, schemes, and equations having maximum widths of 11.3 cm (one-column format) or 23.6 cm (two-column format). Note: if the foregoing

ruler is calibrated in cm. Also note that a standard sheet of paper is only 21.6 cm wide, so all graphics submitted in two-column format must be prepared and printed in landscape mode.

6. Use boldface type for compound numbers but not for atom labels or captions.

7. Label drawings "Reduce to 75% of current size for publication".

Authors using other drawing packages should, as far as possible, modify their program's parameters to reflect the above guidelines.

Supporting Information. Extensive tables, graphs, spectra, calculations, and other material beyond a modest content in the printed paper may be included in the World Wide Web edition of the journal. Three copies of the Supporting Information, one of which is scanner-ready material, should accompany the manuscript for consideration by the editor and reviewers. The material should be described in a paragraph at the end of the paper, using the following format: "Supporting Information Available: Description. This material is available free of charge via the Internet at <http://pubs.acs.org>." Supporting Information should be clear and of high contrast on 8.5 × 11 in. paper. Captions or legends for figures, spectra, etc., must appear directly on the figure. Pages should be consecutively numbered

ADDITIONAL INFORMATION

Reporting Specific Data. For manuscripts in which gas chromatographic methods are used, see "Reporting of Gas Chromatographic Methods", by Morton Beroz and Irwin Hornstein [*J. Agric. Food Chem.* 1973, 21 (1), 7A, located at the back of the issue].

For manuscripts reporting spectral data, see "Guidelines for Presentation of Spectral Data" [*J. Agric. Food Chem.* 1973, 21 (1), 7A, located at the back of the issue] and under the section "Characterization" of the "199 Guidelines for Authors" [*J. Org. Chem.* 1999, 64, 17A].

It is essential that novel compounds, either synthetic or isolated from natural sources, be characterized rigorously and unequivocally. Supporting data normally include physical form, melting point (if solid), UV/vis spectra if appropriate, ^1H and ^{13}C NMR, mass spectral data, and optical rotation (when compounds have chiral centers).

Soils used in research in pesticide and fertilizer metabolism, soil binding, degradation, and other environmental quality factors involving soil-chemical interactions should be described down to the family level according to the soil classification scheme given in *Soil Taxonomy, A Basic System of Soil Classification, Making and Interpreting Soil Surveys* (Agricultural Handbook 436; U.S. Government Printing Office: Washington, DC, 1975). Also give series name if known. The requirement is to allow comparison and extrapolation to other work giving similar soil classifications, published in journals such as the *Journal of Soil Science*, *Soil Science Society of America Journal*, *Journal of Environmental Quality*, and *Geoderma*. If information is unavailable to classify the soils at the desired family level, classification should be described or estimated at least to the great group level in the soil classification system.

Manuscripts reporting on flavor constituents should conform to the recommendations made by the International Organization of the Flavor Industry [for details see the Editorial in the October 1996 issue of *J. Agric. Food Chem.*].

Food Chem. (Vol. 44, No. 10)]. In brief, any identification of a flavoring substance must pass scrutiny of the latest forms of available analytical techniques. In practice, this means that any particular substance must have its identity confirmed by at least two methods, for example, comparison of chromatographic and spectroscopic data (which may include GC, MS, IR, and NMR) with those of an authentic sample (see *Z. Lebensm. Unters. Forsch.* 1991, 192, 530–534). If only one method has been applied (MS data alone or GC relative retention time alone), the identification shall be labeled “tentative”. In addition, authors are encouraged to include at least semiquantitative data on the concentration of an identified component in the original source, for example, foodstuff or plant part. Ranges such as <1 ppb, 1–10 ppb, 10–100 ppb are acceptable.

Electronic Copy of Manuscript. The final accepted version of the manuscript should be submitted as three hard copies and an electronic copy on disk. Submissions not in electronic form may face a delay in publication. Failure to adhere to the following guidelines may prevent all or part of the material supplied on disk from being used in production.

A hardcopy version (three copies) of the manuscript is required for review. The disk should accompany the final accepted version of the manuscript. **The version on the disk must exactly match the final version accepted in hardcopy.**

When preparing a manuscript, use the document mode or its equivalent in the word-processing program, that is, do not save files in “Text Only” (ASCII) mode. If a non-Western version of the word-processing software is used to prepare the manuscript, save the file in rich-text format (RTF). Do not include any page-layout instructions such as placement information for graphics in the file. The text should be left justified, and automatic end-of-line hyphenation should be turned off. Use carriage returns only to end headings and paragraphs, not to break lines of text. Do not insert spaces before punctuation. To ensure expeditious processing of a manuscript, the references should conform to the format described under Literature Cited. Ensure that all characters are correctly represented throughout the manuscript: for example, 1 (one) and l (ell), 0 (zero) and O or o (oh), x (ex) and × (times sign). Check the final copy carefully for consistent notation and correct spelling. The Editorial Office conversion program will faithfully translate any errors to the typeset copy. Check the disk with an up-to-date virus detection program. The presence of a virus may delay the publication of the manuscript. Label the disk with manuscript number and corresponding author name. Provide the platform, version of software used, and filenames on the Diskette Description form.

All of the text (including the title page, abstract, all sections of the body of the paper, figure captions, scheme or chart titles and footnotes, and references) and tabular material should be in one file, with the complete text first followed by the tabular material. It is best to use the fonts “Times” and “Symbol”. Other fonts, particu-

larly those that do not come bundled with the system software, may not translate properly. Ensure that all special characters (Greek characters, math symbols, etc.) are present in the body of the text as characters and not as graphic representations. Consult the documentation for the specific software package being used on how to detect the presence of graphics in the files and replace them with the appropriate text characters.

Tables may be created using a word processor's text mode or table format feature. The table format feature is preferred. Ensure each data entry is in its own table cell. If the text mode is used, separate columns with a single tab and use a line feed (return) at the end of each row.

Procedures for using graphics are currently being refined, and their use in production cannot be guaranteed at this time. If you are able to provide graphics in electronic form, please save each graphic in a separate file in TIFF. Line art should be saved at 600 dots per inch (dpi) and color and grayscales at 300 dpi. Consult the documentation in your graphics application for more information on how to save your files in this format. The filename for each graphic should be descriptive of the content graphic, for example, figure1 for Figure 1 and scheme1 for Scheme 1. Although we cannot guarantee that your electronic files will be used in production, providing the files will help our development efforts.

As additional features become available, these instructions will be updated on the ACS files server. Publication information can be found at <http://pubs.acs.org>; select “Journals & Magazines”, choose a title, and then select “Info for Authors”.

Currently Acceptable Word-Processing Packages. Macintosh: WordPerfect 3.5, Microsoft Word, up to Word 98, and FrameMaker 5.5.

IBM and compatibles: WordPerfect, up to version 8.0, Microsoft Word, up to Word 2000, and FrameMaker 5.5. TeX users should follow the guidelines given at <http://pubs.acs.org/instruct/texguide.html>.

Proofs and Reprint Order Forms. A galley proof, the original and edited manuscripts, and a reprint order form are sent directly to the author who submitted the manuscript. All corrections, revisions, and additions must be entered on the proof and not on the manuscript. Proofs should be checked with utmost care against the manuscript as this is not done when the corrected proof is returned. The corrected proof should be returned as promptly as possible (by airmail from outside North America). No paper is released for publication until the author's corrected proof has been received. Please return the reprint order form purchase order, or check, using the enclosed envelope addressed to Cadmus Professional Communications. Reprints will be shipped within two weeks after the printed journal date. Corresponding authors may obtain 50 free electronic reprints via an Electronic Reprint URL.

These instructions and a copyright status form are printed in the first issue of each volume. Please conform to these instructions when submitting manuscripts.